(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関 国際事務局



(43) 国際公開日 2004 年12 月2 日 (02.12.2004)

PCT

(10) 国際公開番号 WO 2004/104198 A1

(51) 国際特許分類7:

5/00, C12Q 1/06, C12N 7/00

C12N 15/11,

(21) 国際出願番号:

PCT/JP2003/015038

(22) 国際出願日:

2003年11月25日(25.11.2003)

(25) 国際出願の言語:

日本語

(26) 国際公開の言語:

日本語

(30) 優先権データ:

特願2003-148242 特願2003-329115 2003年5月26日(26.05.2003) J. 2003年9月19日(19.09.2003) J.

(71) 出願人 (米国を除く全ての指定国について): 東レ株式会社 (TORAY INDUSTRIES, INC.) [JP/JP]; 〒103-8666 東京都中央区日本橋室町 二丁目 2番1号 Tokyo (JP). 財団法人東京都医学研究機構 (TOKYO METROPOLITAN ORGANIZATION FOR MEDICAL RESEARCH) [JP/JP]; 〒163-8001 東京都新宿区西新宿二丁目8番1号 Tokyo (JP). ヨハネス グーテンベルク ウニベルスィテート マインツ (JOHANNES

GUTENBERG-UNIVERSITAET MAINZ) [DE/DE];

(72) 発明者; および

D-55099 マインツ Mainz (DE).

(75) 発明者/出願人 (米国についてのみ): 脇田 隆字 (WAKITA,Takaji) [JP/JP]; 〒175-0094東京都板橋区成増3-37-1-302 Tokyo (JP). 加藤 孝宜 (KATO,Takanobu) [JP/JP]; 〒467-0004 愛知県 名古屋市瑞穂区松月町1-41-206 Aichi (JP). 伊達朋子 (DATE,Tomoko) [JP/JP]; 〒211-0044 神奈川県川崎市中原区 新城3-13-5-303 Kanagawa (JP).

- (74) 代理人: 平木 祐輔, 外(HIRAKI,Yusuke et al.); 〒 105-0001 東京都港区虎ノ門一丁目17番1号虎ノ門5 森ピル 3階 Tokyo (JP).
- (81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) 指定国 (広域): ARIPO 特許 (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

規則4.17に規定する申立て:

すべての指定国のための不利にならない開示又は新 規性喪失の例外に関する申立て(規則4.17(v))

添付公開書類:

- -- 国際調査報告書
- 不利にならない開示又は新規性喪失の例外に関する 申立て

2文字コード及び他の略語については、定期発行される 各PCTガゼットの巻頭に掲載されている「コードと略語 のガイダンスノート」を参照。

(54) Title: NUCLEIC ACID CONSTRUCT CONTAINING HEPATITIS C VIRUS (HCV) OF GENOCYPE 2a GENOME-ORIGIN NUCLEIC ACID AND CELL HAVING THE NUCLEIC ACID CONSTRUCT TRANSFERRED THEREINTO

(54) 発明の名称: 遺伝子型2aのC型肝炎ウイルス(HCV)ゲノム由来の核酸を含む核酸構築物、及び該核酸構築物 を導入した細胞

(57) Abstract: A replicon RNA comprising a base sequence which contains at least 5'-nontranslational region, base sequences encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein and 3'-nontranslational region on the genomic RNA of hepatitis C virus of genotype 2a.

(57) 要約: 本発明は、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4A タンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする塩基配列と、3'非翻訳領域とを 「少なくとも含む塩基配列からなる、レプリコンRNAに関する。



明細書

遺伝子型2aのC型肝炎ウイルス(HCV)ゲノム由来の核酸を含む核酸構築物、 及び該核酸構築物を導入した細胞

5

技術分野

本発明は、遺伝子型2aのC型肝炎ウイルスのレプリコンRNA、該レプリコンRNAを導入したレプリコン複製細胞、及び該レプリコンRNAの複製効率を増大させる方法に関する。

10

15

20

25

背景技術

C型肝炎ウイルス(Hepatitis C virus、HCV)は、フラビウイルス科に属する、一本鎖の(+)鎖センスRNAをゲノムとするウイルスであり、C型肝炎の原因となることが知られている。近年の研究により、C型肝炎ウイルスは遺伝子型又は血清型により多数の型に分類されることが分かってきた。現在主流であるHCV遺伝子型の分類法である、SimmondsらによるHCV 株の塩基配列を用いた系統解析法ではHCV は遺伝子型1a、遺伝子型1b、遺伝子型2a、遺伝子型2b、遺伝子型3a、遺伝子型3bの6タイプに分類され(Simmonds、P. et al, Hepatology、(1994) 10、p. 1321-1324を参照)、さらにそれらの各タイプがいくつかのサブタイプに分類されている。現在では、HCVの複数の遺伝子型についてゲノム全長の塩基配列が決定されている(特開2002-171978号公報、Choo et al., Science、(1989) 244、p. 359-362、Kato et al., J. Med. Virol., (2001) 64(3) p. 334-339、Okamoto、Het al, J. Gen. Virol., (1992) 73 p. 673-679、及びMori、S. et al, Biochem. Biophis. Res. Commun., (1992) 183、p. 334-342を参照)。

HCVは持続的に感染することにより慢性肝炎を引き起こす。現在、世界的規模で認められる慢性肝炎の主たる原因がHCV持続感染である。実際、持続感染者の50%程度が慢性肝炎を発症し、そのうち約20%の患者が10年~20年を経て肝硬変に移行し、さらにその一部は肝癌といった致死的な病態へと進展する。

C型肝炎に対する現在の主な治療は、インターフェロンー α 、インターフェロンー β 、及びインターフェロンー α とプリンーヌクレオシド誘導体であるリバビリンとの併用療法により行われている。しかしながら、これらの治療を行っても、全治療者の約60%に治療効果が認められるだけであり、効果が出た後に治療を中止すると半分以上の患者が再燃する。インターフェロンの治療効果は、HCVの遺伝子型と関連することが知られており、遺伝子型1bに対しては効果が低く、遺伝子型2aに対してはより効果が高いと言われている(Yoshioka et al., Hepatology, (1992) 16(2): p. 293-299を参照)。

5

10

工業国において罹患率が高く、最終的に深刻な結果を招き、かつ現在は原因治療法が存在しないC型肝炎に対する効果的な治療薬又は予防薬の開発は重要な目標である。そのため、HCV特異的な化学療法、ワクチン療法の発展が切望されている。抗HCV薬開発のターゲットとしては、HCVの複製抑制やHCVの細胞感染の抑制が考えられる。

最近まで、HCVを細胞培養系で増やすこと、培養細胞に感染させることは困難 であり、また、HCVに感染可能かつ実験可能な動物はチンパンジーしかなかった 15 ため、HCVの複製機構や感染機構の研究は困難であった。しかし最近になって、 HCV由来の自律複製能を有するRNAとして、HCVサブゲノムRNAレプリコンが作製さ れたことにより (特開2001-17187号公報、Lohmann et al., Science, (1999) 285, p. 110-113, Blight et al., Science, (2000) 290, p. 1972-1974, Friebe et al., J. Virol., (2001) 75(24): p. 12047-12057、及びIkeda et 20 al., J. Virol., (2002) 76(6): p. 2997-3006を参照)、培養細胞を用いてHCV の複製機構を解析することが可能となった。これらのHCVサブゲノムRNAレプリコ ンは、遺伝子型1bのHCVゲノムRNAの5'非翻訳領域中のHCV IRESの下流に存在する 構造タンパク質を、ネオマイシン耐性遺伝子及びその下流に連結したEMCV IRES によって置換したものである。このRNAレプリコンは、ヒト肝癌細胞Huh7に導入 25 してネオマイシン存在下で培養することにより、Huh7細胞内で自律複製すること が証明された。

しかしながら、このようなHCVの細胞内RNA複製系は、未だ遺伝子型1bのHCVの ゲノムRNAを用いたものしか作製されていない。異なる遺伝子型のHCVではコード

されるウイルスタンパク質にも違いがあることが報告されていることから、遺伝子型1bのHCV由来のサブゲノムRNAレプリコンの解析だけでは、HCVの複製機構を十分に解明することは難しいと考えられる。さらに、インターフェロンの治療効果がHCVの遺伝子型によって異なることから、遺伝子型1bのHCVのサブゲノムRNAレプリコンを含むHCV複製系のみを用いて色々なタイプのHCVに効果を及ぼす抗HCV率を開発することは特に難しいと考えられる。

発明の開示

5

10

20

本明細書は、本願の優先権主張の基礎となる特願2003-148242号及び特願2003-329115号に記載された内容を包含する。

本発明は、未だレプリコンRNAが作製されていない遺伝子型のHCV由来のレプリコンRNAを提供することを目的とする。

本発明者らは、上記課題を解決すべく鋭意研究を行った結果、遺伝子型2aの HCVのレプリコンRNAを作製することに成功した。

15 すなわち、本発明は以下のとおりである。

- [1] 遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする塩基配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNA。このレプリコンRNAは、少なくとも1つの選択マーカー遺伝子又はリポーター遺伝子、及び少なくとも1つのIRES配列をさらに含むことが好ましい。
- [2] 配列番号9又は10のいずれか1つで示される塩基配列からなる5'非翻訳 領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、 IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、
- 25 NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコード する塩基配列と、配列番号11又は12のいずれか1つで示される塩基配列から なる3、非翻訳領域とを含む塩基配列からなる、レプリコンRNA。
 - [3] 遺伝子型2aのC型肝炎ウイルスのゲノムRNAが、配列番号3又は5で示される塩基配列からなるRNAである、上記[1]又は[2]記載のレプリコンRNA。

- [4] 以下の(a)又は(b)のRNAからなるレプリコンRNA。
- (a) 配列番号1又は2で示される塩基配列からなるRNA
- (b) 配列番号1又は2で示される塩基配列において1~10個の塩基が欠失、置換又は付加された塩基配列からなり、かつ、自律複製能を有するRNA
- 5 [5] 上記[1]~[4]記載のいずれかのレプリコンRNAを細胞に導入することにより作製された、レプリコン複製細胞。このレプリコン複製細胞において、レプリコンRNAを導入する細胞は、真核細胞であることが好ましく、ヒト肝由来細胞、ヒト子宮頸由来細胞、又はヒト胎児腎由来細胞であることがより好ましく、Huh7細胞、HepG2細胞、IMY-N9細胞、HeLa細胞、及び293細胞からなる群より選ばれるいずれかの細胞であることがさらに好ましい。
 - [6] C型肝炎ウイルス感染の治療剤若しくは診断剤の製造、又は評価のための、 上記[1]~[4]記載のレプリコンRNA。
 - [7] C型肝炎ウイルス感染の治療剤若しくは診断剤の製造、又は評価のための、 上記[5]記載のレプリコン複製細胞。
- 15 [8] C型肝炎ウイルス感染に対するワクチンの製造のための、上記[1]~[4]記載のレプリコンRNA。
 - [9] C型肝炎ウイルス感染に対するワクチンの製造のための、上記[5]記載のレプリコン複製細胞。
- [10] 上記[5]記載のレプリコン複製細胞からレプリコンRNAを抽出することを含20 む、遺伝子型2aのC型肝炎ウイルスのレプリコンRNAの製造方法。
 - [11] 上記[5]記載のレプリコン複製細胞を培養し、得られる培養物からウイルスタンパク質を取得することを含む、遺伝子型2aのC型肝炎ウイルスのウイルスタンパク質の製造方法。
- [12] 被験物質の存在下で、上記[5]記載のレプリコン複製細胞を培養し、得ら 25 れる培養物中のレプリコンRNAの複製を検出することを含む、C型肝炎ウイルス の複製を促進又は抑制する物質をスクリーニングする方法。
 - [13] 上記[5]記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上行うことを含む、遺伝子型2aのC

型肝炎ウイルスのレプリコンRNAの複製効率を増大させる方法。この方法においては、複製効率の増大が、レプリコン複製細胞に最初に導入されたレプリコンRNAの複製効率と比較して、少なくとも2倍の増大であることがより好ましい。

- [14] 上記[5]記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上行うこと、及び最終的に得られたレプリコン複製細胞から複製レプリコンRNAを取得することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。
- [15] 上記[14]記載の方法によって製造された複製効率が増大したレプリコン RNAについて、レプリコン複製細胞に最初に導入されたレプリコンRNAとの間の塩 基変異又はアミノ酸変異を検出すること、及び複製効率を増大させようとするレ プリコンRNAにその検出された塩基変異又はアミノ酸変異を導入することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する 方法。
- 15 [16] 配列番号1で示される塩基配列上において、以下の(a)~(u):
 - (a) 塩基番号7157の部位におけるAからGへの変異、

5

- (b) 塩基番号4955の部位におけるCからUへの変異、
- (c) 塩基番号4936の部位におけるAからGへの変異、
- (d) 塩基番号5000の部位におけるAからGへの変異、
- 20 (e) 塩基番号 7 2 8 8 の部位におけるAからGへの変異、
 - (f) 塩基番号5901の部位におけるGからUへの変異、
 - (g) 塩基番号6113の部位におけるAからUへの変異、
 - (h) 塩基番号2890の部位におけるAからGへの変異、
 - (i) 塩基番号 6 8 2 6 の部位におけるCからAへの変異、
- 25 (j) 塩基番号 6 8 8 7 の部位におけるCからAへの変異、
 - (k) 塩基番号 6 5 8 0 の部位におけるUからAへの変異、
 - (1) 塩基番号7159の部位におけるUからCへの変異、
 - (m) 塩基番号7230の部位におけるUからAへの変異、
 - (n) 塩基番号 6 9 4 3 の部位におけるCからAへの変異、

- (o) 塩基番号 5 6 8 7 の部位におけるGからAへの変異、
- (p) 塩基番号 6 1 1 0 の部位におけるAからGへの変異、
- (a) 塩基番号5550の部位におけるUからCへの変異、
- (r) 塩基番号7217の部位におけるAからGへの変異、
- 5 (s) 塩基番号3643の部位におけるAからGへの変異、
 - (t) 塩基番号5851の部位におけるGからAへの変異、及び
 - (u) 塩基番号 5 9 1 4 の部位におけるGからAへの変異、からなる群より選択される少なくとも1 つの変異を有する塩基配列からなるレプリコンRNA。
- 本発明により、遺伝子型2aのHCV株に由来するHCV-RNAレプリコンが初めて提供された。本発明に係るレプリコン複製細胞は、遺伝子型2aのHCV由来のRNA及びHCVタンパク質を持続的に産生させるための培養系として用いることができる。さらに本発明に係るレプリコン複製細胞は、HCVの複製及び/又はHCVタンパク質の翻訳に影響を及ぼす各種物質をスクリーニングするための試験系として有用である。

図面の簡単な説明

20

25

図1は、本発明に係るHCV-RNAレプリコンを作製するための鋳型DNAの構築手順を示す概略図である。図1の上段は、pJFH1及びpJCH1のウイルスゲノム挿入領域の構造を示す。図1の下段は、pJFH1及びpJCH1のウイルスゲノム挿入領域の構造を示す。図1の下段は、pJFH1及びpJCH1のウイルスゲノム挿入領域の一部をネオマイシン耐性遺伝子とEMCV IRESを含むDNA断片で置換することにより構築したプラスミドDNA pSGREP-JFH1及びpSGREP-JCH1のウイルスゲノム挿入領域の構造を示す。図中の記号は以下のとおりである。T7: T7 RNAプロモーター、G: 挿入したJFH-1又はJCH-1由来DNAの5'端の上流かつT7 RNAプロモーター配列の3'端の下流に挿入したdGTP、5'NTR:5'非翻訳領域、Core:コアタンパク質、3'NTR:3'非翻訳領域。E1、E2:エンベロープタンパク質。NS2、NS3、NS4A、NS4B、NS5A、NS5B:非構造タンパク質。Age I、Cla I、Xba I: 制限酵素Age I、Cla I及びXba Iの切断部位。GDD: NS5Bタンパク質の活性中心に相当するアミノ酸モチーフGDDの位置。neo:ネオマイ

シン耐性遺伝子、EMCV IRES: EMCV IRES(脳心筋炎ウイルスの内部リボソーム結合部位)。

図2A~Fは、rSGREP-JFH1の塩基配列を示す。

図3A~Fは、rSGREP-JCH1の塩基配列を示す。

5 図4は、rSGREP-JFH1、rSGREP-JFH1/GND及びrSGREP-JFH1/dGDDをそれぞれトランスフェクションしたHuh7細胞のコロニー形成を示した写真である。トランスフェクションしたRNAの量は、上段の3つはいずれも100 ng、下段の3つはいずれも300 ngである。各培養ディッシュには、1.0mg/mlのG418を添加した。

図 5 は、rSGREP-JFH1及びrSGREP-JCH1をそれぞれトランスフェクションした Huh7細胞の、培地中のG418の濃度が0.5mg/m1である場合のコロニー形成を示した 写真である。トランスフェクションしたRNAの量は、いずれも100mgである。

図6は、トランスフェクション細胞のコロニー形成能に対する、Mung Bean Nuclease処理の影響を示した写真である。トランスフェクションしたrSGREP-JFH1 RNAの量は、いずれも100 ngである。いずれも培地中のG418の濃度は

15 1.0mg/mlである。

20

図7は、rSGREP-JFH1をトランスフェクションして樹立したレプリコン複製細胞クローン由来のトータル細胞性RNAを新たなHuh7細胞に再トランスフェクションした場合に示されるコロニー形成を示す写真である。左側写真は、レプリコン複製細胞クローン6由来のトータル細胞性RNAを用いて96コロニーの形成がみとめられた結果、右側写真:プールクローン由来のトータル細胞性RNAを用いて77コロニーの形成がみとめられた結果を示す。いずれもレプリコンRNAを1×10⁷コピー含む量を再トランスフェクションした。

図8は、rSGREP-JFH1をトランスフェクションして樹立したレプリコン複製細胞クローン由来のトータル細胞性RNAを新たなHuh7細胞に再トランスフェクションして得た細胞クローン由来のトータルRNAに対して、rSGREP-JFH1特異的プローブを用いてノーザンブロット法による検出を行った結果を示す写真である。レーンの説明は以下のとおりである。10⁸: Huh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の8乗コピー加えたサンプル、10⁷: Huh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の7乗コ

ピー加えたサンプル、Huh7: トランスフェクションしていないHuh7細胞から抽出したトータルRNA、プールクローン: プールクローンから抽出したトータルRNA、1~11: 細胞クローン1~11のそれぞれから抽出したトータルRNA。「レプリコンRNA」は、rSGREP-JFH1の分子量サイズを示すマーカー、「28S」は4.5kbの分子量サイズを示すリボソームRNAマーカー、「18S」は1.9kbの分子量サイズを示すリボソームRNAマーカーの泳動位置を示す。

5

10

図9は、rSGREP-JFH1又はrSGREP-JCH1由来複製レプリコンRNAが再トランスフェクションされた細胞クローンにおける、ネオマイシン耐性遺伝子の宿主細胞のゲノムDNAへの組み込みの有無を示す写真である。左側写真のレーンの説明は以下のとおりである。M: DNA分子量マーカー、1~8: rSGREP-JFH1由来細胞クローン1~8、N: トランスフェクションしていないHuh7細胞、P: 陽性対照(ネオマイシン耐性遺伝子のPCR増幅産物)。一方、右側写真のレーンの説明は以下のとおりである。M: DNA分子量マーカー、1~6: rSGREP-JCH1由来細胞クローン1~6。

図10は、rSGREP-JFH1又はrSGREP-JCH1由来複製レプリコンRNAが再トランスフェクションされた細胞クローンにおいて、発現されたNS3タンパク質の検出結果を示す写真である。左側写真のレーン1~8: rSGREP-JFH1由来細胞クローン1~8。右側写真のレーン1~6: rSGREP-JCH1由来細胞クローン1~6。右側写真のレーンのP: NS3タンパク質(陽性対照)、N: トランスフェクションしていないHuh7細胞から抽出したタンパク質(陰性対照)。

図11は、rSGREP-JFH1由来の複製レプリコンRNAのHuh7細胞への再トランスフェクションを経て樹立した21の細胞クローンから取得したレプリコンRNA中の塩基変異の位置を示す図である。変異の位置は、表4に示す塩基番号を付記した縦線で示している。太い縦線は非同義置換、細い縦線は同義置換を表す。

25 図12は、A: HepG2細胞、B: IMY-N9細胞、C: 293細胞、又はD: Hela細胞を用いたrSGREP-JFH1トランスフェクションの結果を示す写真である。各培養ディッシュには、0.8 mg/mlのG418を添加した。

図13は、レプリコン複製細胞クローンのノーザンブロッティングの結果を示す写真である。レーン1: HepG2 (陰性対照)、レーン2: 合成RNA 10⁸コピー、

レーン 3: 合成RNA 10⁷コピー、レーン 4: Hep-IH-1 (HepG2細胞由来)、レーン 5: Hep-IH-3 (HepG2細胞由来)、レーン 6: Hep-IH-5 (HepG2細胞由来)、レーン 7: Hep-IH-11 (HepG2細胞由来)、レーン 8: Hep-IH-13 (HepG2細胞由来)、レーン 9: IMY-IH-3 (IMY-N9細胞由来)、レーン 10: IMY-IH-4 (IMY-N9細胞由来)、レーン 11: IMY-IH-7 (IMY-N9細胞由来)、レーン 12: IMY-IH-10 (IMY-N9細胞由来)、レーン 13: 293-IHをトランスフェクトした細胞プール (293細胞由来)、レーン 14: HeLa-IH-9 (HeLa細胞由来)、レーン 15: HeLa-IH-12 (HeLa細胞由来)、レーン 16: HeLa-IH-13 (HeLa細胞由来)、レーン 17: HeLa (陰性対照)。

10 図14は、ネオマイシン耐性遺伝子のゲノムDNAへの組み込みを確認するため の電気泳動の結果を示す写真である。HepG2レプリコン細胞(上段)及びIMY-N9 レプリコン細胞(下段)のそれぞれのゲノムDNAについて、PCR解析によりネオマ イシン耐性遺伝子を検出した。M: DNAサイズマーカー、P: 陽性対照、H: HepG2細胞、I: IMY-N9細胞、▶: PCR産物。

15 図15は、レプリコン複製細胞クローン由来のタンパク質NS3及びNS5aをウエスタンブロット法により解析した結果を示す写真である。

発明を実施するための最良の形態 以下、本発明を詳細に説明する。

20 1. 本発明に係るHCV由来のレプリコンRNA

5

25

C型肝炎ウイルス (HCV) のゲノムは、約9600ヌクレオチドかちなる (+)鎖の一本鎖RNAである。このゲノムRNAは、5'非翻訳領域 (5'NTR又は5'UTRとも表記する)、構造領域と非構造領域とから構成される翻訳領域、及び3'非翻訳領域 (3'NTR又は3'UTRとも表記する)からなる。その構造領域にはHCVの構造タンパク質がコードされており、非構造領域には複数の非構造タンパク質がコードされている。

このようなHCVの構造タンパク質と非構造タンパク質は、翻訳領域から一続きのポリプロテインとして翻訳された後、プロテアーゼによって限定分解を受けて構造タンパク質 (Core、E1、及びE2) と非構造タンパク質 (NS2、NS3、NS4A、

NS4B、NS5A、及びNS5B)とが各タンパク質として遊離することにより、生成される。これらの構造タンパク質及び非構造タンパク質(すなわち、HCVのウイルスタンパク質)のうち、Coreはコアタンパク質であり、E1及びE2はエンベロープタンパク質であり、非構造タンパク質(NS2、NS3、NS4A、NS4B、NS5A、及びNS5B)はウイルス自身の複製に関与するタンパク質である。NS2はメタロプロテアーゼ活性、NS3はセリンプロテアーゼ活性(N末端側の3分の1)とヘリカーゼ活性(C末端側の3分の2)を有することが知られている。またNS4AはNS3のプロテアーゼ活性に対するコファクターであり、NS5BはRNA依存RNAポリメラーゼ活性を有することも報告されている。そして、遺伝子型2aのHCVのゲノムも同様の遺伝子構造を有することがすでに報告されている(特開2002-171978号公報を参照)。

5

10

15

20

25

本発明者らは、このような遺伝子型2aのHCVゲノムを用いて、自律的に複製することが可能なRNAを構築した。すなわち本発明のHCV由来のレプリコンRNAは、遺伝子型2aのHCVゲノムの全体又は部分RNAを含む自律複製能を有するRNA構築物である。

本明細書では、自律複製能を有しておりHCVウイルスゲノムを改変して作製されたRNAを、「レプリコンRNA」又は「RNAレプリコン」と呼び、遺伝子型2aのHCVから人為的に作製される自律複製能を有するRNAを、遺伝子型2aのHCV由来のレプリコンRNAと称する。本明細書においてHCV由来のレプリコンRNAは、HCV-RNAレプリコンとも称する。

本発明において、「遺伝子型2aのC型肝炎ウイルス」「遺伝子型2aのHCV」とは、Simmondsらによる国際分類に従って遺伝子型2aと同定されるC型肝炎ウイルスを意味する。本発明における「遺伝子型2aのC型肝炎ウイルス」「遺伝子型2aのHCV」には、天然由来のHCVゲノムRNAを有するウイルスだけでなく、天然由来のHCVゲノム配列に人為的な改変を加えたゲノムRNAを有するウイルスも包含する。遺伝子型2aのHCVの具体例としては、JFH-1株及びJCH-1株(特開2002-171978号公報を参照)等のウイルスが挙げられる。

さらに「遺伝子型2aのC型肝炎ウイルスのゲノムRNA」とは、遺伝子型2aのC型肝炎ウイルスの一本鎖の(+)鎖センスRNAからなるゲノムの全領域にわたる塩基

配列を有するRNAを意味する。限定するものではないが、遺伝子型2aのC型肝炎ウイルスのゲノムRNAは、好ましくは配列番号3又は5で示される塩基配列からなるRNAである。

5

10

15

20

25

本願明細書において、「5'非翻訳領域(5'NTR又は5'UTR)」、「NS3タンパク 質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコ ードする配列」、「Coreタンパク質をコードする配列(Core領域又はC領域)」、 「E1タンパク質をコードする配列(E1領域)」、「E2タンパク質をコードする配 列 (E2領域) 」、「N2タンパク質をコードする配列 (NS2領域) 」、「NS3タンパ ク質をコードする配列 (NS3領域)」、「NS4Aタンパク質をコードする配列 (NS4A領域)」、「NS4Bタンパク質をコードする配列 (NS4B領域)」、「NS5Aタン パク質をコードする配列 (NS5A領域)」、「NS5Bタンパク質をコードする配列 (NS5B領域)」、及び「3'非翻訳領域(3'NTR又は3'UTR)」、並びにその他の特定 の領域若しくは部位は、遺伝子型2aのHCVであるJFH-1株のゲノム全領域をコード する全長cDNA (JFH-1クローン) の塩基配列(配列番号3)を基準として、定め るものとする。配列番号3の塩基配列は、国際DNAデータバンク(DDBJ/EMBL/GenBank) からアクセッション番号AB047639により取得可能である。 具体的には、配列番号3で示される塩基配列に対して特定のHCVのRNA配列をアラ インメントしたときに、配列番号3で示される塩基配列上の塩基番号1~340に アラインメントされる配列がそのRNAの「5'非翻訳領域」、同塩基番号3431~ 9442にアラインメントされる配列が「NS3タンパク質、NS4Aタンパク質、NS4Bタ ンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」、同塩基番号 3431~5323にアラインメントされる配列が「NS3タンパク質をコードする配列」、 塩基番号5324~5485にアラインメントされる配列が「NS4Aタンパク質をコードす る配列」、同塩基番号5486~6268にアラインメントされる配列が「NS4Bタンパク 質をコードする配列」、同塩基番号6269~7666にアラインメントされる配列が「 NS5Aタンパク質をコードする配列」、塩基番号7667~9442にアラインメントされ る配列が「NS5Bタンパク質をコードする配列」、同塩基番号9443~9678にアライ ンメントされる配列が「3'非翻訳領域」である。また、この場合「アラインメン ト」される配列にはギャップ、付加、欠失、置換等が存在していてもよい。さら

に上記の「特定のHCV」は、限定するものではないが、JFH-1株若しくはJCH-1株 又はそれらの誘導体であるウイルス株を包含する。

本発明に係るHCV RNA-レプリコンの一つの実施形態は、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNAである。このレプリコンRNAは、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子、及び少なくとも1つのIRES配列をさらに含んでもよい。さらにこのレプリコンRNAは、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、NS3、NS4A、NS4B、NS5A及びNS5Bタンパク質以外のウイルスタンパク質をコードする配列を、含んでもよい。

5

10

15

20

25

本発明に係るHCV RNA-レプリコンの別の好適な実施形態は、配列番号9又は10で示される塩基配列からなる5'非翻訳領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列と、配列番号11又は12で示される塩基配列からなる3'非翻訳領域とを含む塩基配列からなる、レプリコンRNAである。ここで配列番号9及び10で示される塩基配列は、それぞれ、本発明に係るレプリコンRNAであるrSGREP-JFH1(配列番号1)及びrSGREP-JCH1(配列番号2)の5'非翻訳領域の配列である。また配列番号11及び12で示される塩基配列は、それぞれ、本発明に係るレプリコンRNAであるrSGREP-JFH1(配列番号1)及びrSGREP-JCH1(配列番号1)及びrSGREP-JCH1(配列番号2)の3'非翻訳領域の配列である。

本発明に係るHCV RNA-レプリコンのさらに好ましい1つの実施形態は、配列番号1又は2で示される塩基配列からなるRNAからなるレプリコンRNAである。さらに、この配列番号1又は2で示される塩基配列において、1~50個、1~30個、1~10個、1~6個、1~数個(2~5個)の塩基が欠失、置換又は付加された塩基配列からなるレプリコンRNAであって、かつ、自律複製能を有するRNAも、好適な実施形態として本発明の範囲に含まれる。本発明において「自律複製能を有する」とは、レプリコンRNAを細胞中に導入したときに、そのレプリコンRNAが細胞内でそ

• 6

のレプリコンRNA自身の全長を複製することができることを意味する。限定するものではないが、この自律複製能は、例えば、レプリコンRNAをHuh7細胞中にトランスフェクションし、Huh7細胞を培養し、得られる培養物中の細胞から抽出したRNAについて、導入したレプリコンRNAを特異的に検出可能なプローブを用いたノーザンブロットハイブリダイゼーションを行ってレプリコンRNAの存在を検出することにより、確認することができる。自律複製能を確認するための具体的な操作は、本明細書の実施例に記載されたコロニー形成能の測定、HCVタンパク質の発現確認、レプリコンRNAの検出等の記載に従って行うことができる。

5

10

15

20

25

本発明において「選択マーカー遺伝子」とは、その遺伝子が発現された細胞だけが選択されるような選択性を細胞に付与することができる遺伝子を意味する。選択マーカー遺伝子の一般的な例としては抗生物質耐性遺伝子が挙げられる。本発明において好適な選択マーカー遺伝子の例としては、ネオマイシン耐性遺伝子、チミジンキナーゼ遺伝子、カナマイシン耐性遺伝子、ピリチアミン耐性遺伝子、アデニリルトランスフェラーゼ遺伝子、ゼオシン耐性遺伝子、ピューロマイシン耐性遺伝子等が挙げられるが、ネオマイシン耐性遺伝子、チミジンキナーゼ遺伝子が好ましく、ネオマイシン耐性遺伝子がさらに好ましい。但し本発明における選択マーカー遺伝子はこれらに限定されるものではない。

また本発明において「リポーター遺伝子」とは、その遺伝子発現の指標となる遺伝子産物をコードするマーカー遺伝子を意味する。リポーター遺伝子の一般的な例としては、発光反応や呈色反応を触媒する酵素の構造遺伝子が挙げられる。本発明において好適なリポーター遺伝子の例としては、トランスポプンTn9由来のクロラムフェニコールアセチルトランスフェラーゼ遺伝子、大腸菌由来のβグルクロニダーゼ若しくはβガラクトシダーゼ遺伝子、ルシフェラーゼ遺伝子、線色蛍光タンパク質遺伝子、クラゲ由来のイクリオン遺伝子、分泌型胎盤アルカリフォスファターゼ(SEAP)遺伝子等が挙げられる。但し本発明におけるリポーター遺伝子はこれらに限定されるものではない。

上記の選択マーカー遺伝子及びリポーター遺伝子は、レプリコンRNA中にどちらか一方のみが含まれていてもよいし、両方が含まれていてもよい。

本発明における「IRES配列」とは、RNAの内部にリボソームを結合させて翻訳

を開始させることが可能な内部リボゾーム結合部位を意味する。本発明における IRES配列の好適な例としては、以下に限定するものではないがEMCV IRES (脳心筋炎ウイルスの内部リボゾーム結合部位)、FMDV IRES、HCV IRES等が挙げられるが、EMCV IRES及びHCV IRESがより好ましく、EMCV IRESが最も好ましい。

5

10

15

20

25

さらに本発明に係るレプリコンRNAは、他のHCV株又は他の遺伝子型のHCVのゲノムRNA上の配列を含んでもよい。例えば、遺伝子型1bのHCVゲノムの断片を含んでもよい。他のHCV株としては、例えばHCV-1、HCV-H、HC-J1、HCT-18、H77、DK-7、US11、S14、HCT23、HCV-Th、DR1、DR4、HCT27、S18、SW1、DK9、H90、TD-6E1、S9、HCV-BK、T10、DK1、HC-J4、HCV-J、HK3、HK8、HK5、HCV-G3、IND5、IND8、P10、D1、D3、SW2、T3、S45、SA10、US6、HCV-JK1、HCV-JK4、HCV-JK3、HCV-JK2、HCV-JT、HC-J2、HCV-T、HK4、HC-G9、Z1、Bi,S.I.、Cho,J.M.、HCV-J6、T4、T9、US10、HC-J5、T2、HC-J7、DK11、SW3、DK8、T8、HC-J8、S83、HK2、HC-J6、HC-J8、BEBE1、HCV-J6、HCV-J8、HD10-2、BR36-9、S52、S54、S2、BR33-1、HK10、DK12、HCV-TR、BA-1、BA-2、DK13、Z1、Z4、Z6、Z7、HK2、SA1、SA4、SA5、SA7、SA13、SA6、NZL1、SA30、EG-13、HCV-K3a/650、ED43、EUH1480、EUHK2、Th580、VN235、VN405、VN004、JK049、JK046、JFH-1、JCH-1、JCH-2、JCH-3、JCH-4、

本発明に係るレプリコンRNAは、好ましくは、最も5'側に遺伝子型2aのHCVのゲノムRNA上の5'非翻訳領域を、最も3'側に遺伝子型2aのHCVのゲノムRNA上の3'非翻訳領域を有する。選択マーカー遺伝子又はリポーター遺伝子は、IRES配列の上流に連結されてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の上流又は下流に連結されてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質、NS5Aタンパク質、NS5Bタンパク質をコードする配列」の中に挿入されてもよい。

JCH-5、JCH-6、J6CF、H77等が挙げられるが、これらに限定するものではない。

本発明に係るレプリコンRNAは、より好ましくは、最も5'側に遺伝子型2aのHCVのゲノムRNA上の5'非翻訳領域を有し、それよりも下流に選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」とをこの順番で有し、さらに最も3'側に遺伝子型2aのHCVのゲノムRNA上の3'非翻訳領域

を有する。

5

10

15

20

25

本発明に係るレプリコンRNAには、上記したような配列の他に、レプリコンRNAを導入する細胞内で発現させたい任意の外来遺伝子を含むRNAを含んでもよい。外来遺伝子は、5'非翻訳領域の下流に連結してもよいし、選択マーカー遺伝子若しくはリポーター遺伝子の上流又は下流に連結させてもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Aタンパク質、NS4Aタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の上流又は下流に連結してもよいし、「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」の中に挿入してもよい。外来遺伝子を含むレプリコンRNAは、導入された細胞内で翻訳される際に、該外来遺伝子にコードされたタンパク質を発現することができる。従って外来遺伝子を含むレプリコンRNAは、遺伝子治療などの、特定の遺伝子産物を細胞内で生成させることを目的とする場合にも、好適に使用することができる。

また本発明に係るレプリコンRNAには、さらにリボザイムを含んでいてもよい。 リボザイムは、5'側のレプリコンRNA中の選択マーカー遺伝子、リポーター遺伝 子又は外来遺伝子と、それより3'側のIRES配列及び「NS3タンパク質、NS4Aタン パク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列 」とを連結するように挿入し、リボザイムの自己切断活性により両者が切断され て分離するようにすることができる。

本発明に係るレプリコンRNAにおいては、上述したような選択マーカー遺伝子、リポーター遺伝子、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のウイルスタンパク質をコードする配列、遺伝子型2a以外のHCVのウイルスタンパク質コード配列、及び外来遺伝子等が、レプリコンRNAから正しい読み枠で翻訳されるように連結される。それらの配列のうちでタンパク質をコードする配列は、遺伝子型2aのC型肝炎ウイルスの「NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする配列」から翻訳されるポリタンパク質とともに融合タンパク質として発現させた後で、プロテアーゼによって各タンパク質へと分離するように、プロテアーゼ切断部位等を介して互いに連結させてもよい。

2. 本発明に係るレプリコンRNAの作製

5

10

15

20

25

本発明に係るHCV RNA-レプリコンは、当業者に公知である任意の遺伝子工学的手法を用いて作製することができる。限定するものではないが、HCV RNA-レプリコンは、例えば以下のような方法で作製することができる。

まず、遺伝子型2aのC型肝炎ウイルスのゲノムRNAの全領域に対応するDNAを、 常法によりRNAプロモーターの下流に連結し、DNAクローンを作製する。ここで、 「RNAに対応するDNA」とは、当該RNAの塩基配列のU(ウラシル)をT(チミン) に置き換えた塩基配列を有するDNAを意味する。前記RNAプロモーターは、プラス ミドクローン中に含まれるものであることが好ましい。RNAプロモーターとして は、限定するものではないが、T7 RNAプロモーターが特に好ましい。

次に、作製したDNAクローンについて、例えば、5'非翻訳領域の下流に位置する構造領域(Core配列、E1配列、E2配列)及びNS2タンパク質をコードする配列を、選択マーカー遺伝子若しくはレポーター遺伝子とその下流に連結したIRES配列とを含むDNA断片によって置換する。この置換においては、構造領域以外の部分、例えば5'非翻訳領域の3'末端側の断片及びNS3タンパク質をコードする配列の一部が、別の遺伝子型のHCVに由来する配列に置換されてもよい。

次いで、その置換したDNAクローンを鋳型として、RNAポリメラーゼによりRNAを合成する。RNA合成は、5'非翻訳領域及びIRES配列から、常法により開始させることができる。鋳型DNAがプラスミドクローンの場合には、そのプラスミドクローンから、RNAプロモーターの下流に連結された上記DNA領域を制限酵素によって切り出して、そのDNA断片を鋳型として用いてRNAを合成することもできる。なお、好ましくは合成されるRNAの3'末端がウイルスゲノムRNAの3'非翻訳領域と一致し、他の配列が付加されたり削除されたりしないことが好ましい。このようにして合成されたRNAが、本発明に係るレプリコンRNAである。

3. 遺伝子型2aのHCVのレプリコンRNAを導入したレプリコン複製細胞の作製

上記のようにして作製されるレプリコンRNAを、レプリコンRNAを複製させるべき細胞に導入することにより、レプリコンRNAが持続的に複製されている細胞を

得ることができる。本明細書では、レプリコンRNAが持続的に増幅されている細胞を「レプリコン複製細胞」と称する。

レプリコンRNAを導入する細胞としては、継代培養することが可能な細胞であれば任意の細胞を用いることができるが、真核細胞であることが好ましく、ヒト肝由来細胞、ヒト子宮頸由来細胞、又はヒト胎児腎由来細胞であることがより好ましく、Huh7細胞、HepG2細胞、IMY-N9細胞、HeLa細胞、及び293細胞からなる群より選ばれるいずれかの細胞であることがさらに好ましい。これらの細胞は、市。販のものを利用してもよいし、細胞寄託機関から入手して使用してもよいし、任意の細胞(例えば癌細胞又は幹細胞)から株化した細胞を使用してもよい。

前記細胞は、ワクチン製造のようにHCVタンパクの大量製造を目的とする場合には、大量培養が可能な細胞を用いることが望ましい。そのような観点からは、Huh7細胞以外の細胞であることが好ましい。

10

15

20

25

レプリコンRNAの細胞内への導入は、当業者には公知の任意の技術を使用して行うことができる。そのような導入法としては、例えば、エレクトロポレーション、パーティクルガン法、リポフェクション法、リン酸カルシウム法、マイクロインジェクション法、DEAEセファロース法等が挙げられるが、エレクトロポレーションによる方法が特に好ましい。

レプリコンRNAは、目的のレプリコンRNAを単独で導入してもよいし、他の核酸と混合させたものを導入してもよい。導入するRNA量を一定にしてレプリコンRNAの量を変化させたい場合には、目的のレプリコンRNAを、導入する細胞から抽出したトータル細胞性RNAと混合して、細胞内導入に用いればよい。細胞内導入に用いるレプリコンRNAの量は、使用する導入法に応じて決めればよいが、好ましくは1ピコグラム~100マイクログラム、より好ましくは10ピコグラム~10マイクログラムの量を使用する。

細胞内導入のために選択マーカー遺伝子又はリポーター遺伝子を含有するレプリコンRNAを用いる場合には、そのレプリコンRNAが導入され持続的に複製している細胞を、選択マーカー遺伝子又はリポーター遺伝子の発現を利用して、選択することができる。具体的には、例えば、そのようなレプリコンRNAの細胞内導入処理を施した細胞を、選択マーカー遺伝子又はリポーター遺伝子の発現により選

択可能となる培地において培養すればよい。一例として、レプリコンRNAにネオ マイシン耐性遺伝子が選択マーカー遺伝子として含まれる場合には、そのレプリ コンRNAを用いて細胞内導入処理した細胞を培養ディッシュに播種し、16~24時 間培養した後に、培養ディッシュにG418(ネオマイシン)を0.05ミリグラム/ミ リリットル~3.0ミリグラム/ミリリットルの濃度で添加し、その後、週に2回培 養液を交換しながら培養を継続し、播種時から好ましくは10日間~40日間、より 好ましくは14日間~28日間培養した後にクリスタルバイオレットで生存細胞を染 色することにより、レプリコンRNAが導入され持続的に複製されている細胞を、

5

15

20

25

形成されたコロニーからは、常法により生存細胞をクローン化し、培養を継続 10 することにより、細胞をクローン化することができる。このようにして得られる 目的のレプリコンRNAが持続的に複製されている細胞クローンを、本明細書では 「レプリコン複製細胞クローン」と称する。

樹立した細胞クローンについては、導入されたレプリコンRNAから該細胞クロ ーン中で複製されているレプリコンRNAの検出、導入されたレプリコンRNA中の選 択マーカー遺伝子又はリポーター遺伝子の宿主ゲノムDNAへの組み込みの有無の 確認、及びHCVタンパク質の発現の確認を行って、実際に目的のレプリコンRNAが 持続的に複製されていることを確認することが好ましい。

導入されたレプリコンRNAから該細胞クローン中で複製されたレプリコンRNA(本明細書中では、以下便宜的に、「複製レプリコンRNA」と称する)の検出は、 当業者には公知の任意のRNA検出法に従って行えばよいが、例えば、細胞クロー ンから抽出したトータルRNAについて、導入されたレプリコンRNAに対して特異的 なDNA断片をプローブとして用いるノーザンハイブリダイゼーション法を実施す ることにより検出することができる。

また導入されたレプリコンRNA中の選択マーカー遺伝子又はリポーター遺伝子 の宿主ゲノムDNAへの組み込みの有無の確認は、限定するものではないが、例え ば、細胞クローンから抽出した宿主ゲノムDNAについて該選択マーカー遺伝子又 はリポーター遺伝子の少なくとも一部を増幅するPCRを行い、その増幅産物の有 無を確認することによって行うことができる。増幅産物が確認された細胞クロー

ンでは、宿主ゲノム中に選択マーカー遺伝子又はリポーター遺伝子が組み込まれていると判断されることから、レプリコンRNA自体は該細胞内で持続的に複製されていない可能性がある。この場合、レプリコンRNAが持続的に複製されているか否かを、次に示すHCVタンパク質の発現の確認実験によって、確認することができる。

5

10

15

20

25

することができる。

HCVタンパク質の発現の確認は、例えば、導入されたレプリコンRNAから発現されるべきHCVタンパク質に対する抗体を、細胞クローンから抽出したタンパク質と反応させることによって行うことができる。この方法は、当業者には公知の任意のタンパク質検出法によって行うことができるが、具体的には例えば、細胞クローンから抽出したタンパク質試料をニトロセルロース膜にブロッティングし、それに対して抗HCVタンパク質抗体(例えば、抗NS3特異的抗体、又はC型肝炎患者から採取した抗血清)を反応させ、さらにその抗HCVタンパク質抗体を検出することによって行うことができる。細胞クローンから抽出したタンパク質中からHCVタンパク質が検出されれば、その細胞クローンは、HCV由来のレプリコンRNAが持続的に複製してHCVタンパク質を発現しているものと判断することができる。

以上のようにして、目的のレプリコンRNAを持続的に複製していることが確認された細胞クローン(レプリコン複製細胞クローン)を得ることができる。また本発明においては、このレプリコン複製細胞から、例えばRNAを抽出しその中からレプリコンRNAを電気泳動法により分離する等の当業者には公知の任意の方法により、レプリコンRNAを取得することができる。本発明はそのようなレプリコンRNAの製造方法にも関する。さらに本発明に係るレプリコン複製細胞は、HCVタンパク質を製造するために好適に使用することができる。レプリコン複製細胞からのHCVタンパク質の取得は、当業者であれば常法に従って行うことができる。具体的には例えば、レプリコン複製細胞を培養し、得られる培養物(培養細胞及び培養培地を含む)から常法によりタンパク質を採取し、さらにそのタンパク質から、抗HCVタンパク質抗体を用いた検出等によりウイルスタンパク質を選択的

また本発明に係るレプリコン複製細胞が、外来遺伝子を含有するレプリコン

に得ることにより、遺伝子型2aのC型肝炎ウイルスのウイルスタンパク質を製造

RNAを持続的に複製している場合には、そのレプリコン複製細胞を用いて外来遺伝子にコードされるタンパク質を発現させて取得することができる。具体的には例えば、レプリコン複製細胞を培養し、得られる培養物(培養細胞及び培養培地を含む)から常法によりタンパク質を採取し、さらにそのタンパク質から、目的のタンパク質に対する抗体を用いた検出等によりタンパク質を選択的に得ることにより、外来遺伝子にコードされたタンパク質を取得することができる。

5

10

15

20

25

4. 遺伝子型2aのHCVのレプリコンRNAへの複製効率を増大させる突然変異の導入本発明に係るレプリコン複製細胞において複製され生成されたレプリコンRNA (複製レプリコンRNA) には、複製効率を向上させる突然変異が頻繁に生ずる。このような突然変異は適合変異であると思われる。

本発明では、このことを利用して、本発明に係るレプリコンRNAに複製効率を向上させる突然変異の導入を促進することができる。

具体的には、第1のレプリコン複製細胞(好ましくは本発明に係るレプリコンRNAを導入したレプリコン複製細胞)から、第1の複製レプリコンRNAを抽出等により取得し、第1の複製レプリコンRNAをさらに別の細胞に再導入して第2のレプリコン複製細胞を作製するという工程を、1回以上、好ましくは1~10回、より好ましくは1~5回、さらに好ましくは1~2回反復的に行うことにより、レプリコン複製細胞中で、レプリコンRNAに複製効率を増大させる突然変異を高頻度に導入することができる。

複製レプリコンRNAを再導入する細胞としては、任意の細胞を用いることができるが、最初にレプリコンRNAを導入した細胞と同じ生物種由来の細胞であることが好ましく、最初にレプリコンRNAを導入した細胞と同じ生物種由来の同じ組織由来の細胞であることが好ましく、最初にレプリコンRNAを導入した細胞と同じ細胞株の細胞であることがさらに好ましい。

従って本発明では、上記の方法を用いて、複製効率を増大させる突然変異を導入したレプリコンRNAを製造することができる。すなわち、まず第1のレプリコン複製細胞(好ましくは本発明に係るレプリコンRNAを導入したレプリコン複製細胞)から、第1の複製レプリコンRNAを抽出等により取得し、さらにこの第1

の複製レプリコンRNAをさらに別の細胞に再導入して第2のレプリコン複製細胞を作製する工程を、1回以上、好ましくは1~10回、より好ましくは1~5回、さらに好ましくは1~2回反復的に行った後、この反復工程の最後に得られる最終的なレプリコン複製細胞から、複製レプリコンRNAを抽出等によって取得することにより、複製効率が増大したレプリコンRNAを製造することができる。

本発明では、以上のような方法により、レプリコンRNAの複製効率を少なくとも2倍、好ましくは10~100倍、より好ましくは100~10000倍に増大させることができる。

5

15

20

25

このような方法により製造した複製効率が増大したレプリコンRNAについては、 逆転写PCRによってcDNAを得てそれを塩基配列決定に供するなどの公知の方法に より、塩基配列を決定することが好ましい。さらに、決定された塩基配列又はそ れにコードされるアミノ酸配列を、最初に細胞に導入されたレプリコンRNAの塩 基配列と比較することにより、適合変異を同定することができる。複製効率を増 大させる適合変異としては、特に、レプリコンRNAにコードされたウイルスタン パク質のアミノ酸を変異させる非同義置換が好ましい。

また本発明は、そのようにして同定した適合変異を、複製効率を増大させようとするレプリコンRNAに常法により導入することによって、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造することができる方法も提供する。

以上のようにして製造された複製効率が増大したレプリコンRNAは、その方法 に使用した細胞中においてレプリコンRNAを大量に製造するために使用すること ができる。

本発明に係るレプリコンRNAの複製効率は、当業者に公知の方法により決定することができるが、例えば次のような方法に従って決定すればよい。たとえば Huh7細胞に0.0001、0.0003、0.001、0.003、0.01、0.03、0.1、0.3、1.0マイクログラムの量のレプリコンRNAをトランスフェクションして、前述の実験手法と同様の方法でG418による選択培養を21日間行った後にコロニー形成数(コロニー数)を測定する。導入したRNA量とコロニー形成数とを比較して容量依存的にコロニー形成が増加するレプリコンRNA導入量の範囲を決定し、その範囲内での

コロニー形成数を、導入したRNA量で除算して得られる値を、1マイクログラム あたりのコロニー形成率とする。この計算式は、以下のとおりである。

コロニー形成率 [(Colony forming Unit; CFU)/マイクログラム] = コロニー形成数 [個]/導入したRNA量 [マイクログラム]

こうして算出されたコロニー形成率を、導入したレプリコンRNAの複製効率を示す値とする。すなわち、コロニー形成率が高いほど、そのレプリコンRNAの複製効率は高い。またレプリコンRNAの複製効率は、形成されたコロニー1個あたりの導入したレプリコンRNAのコピー数で示されるコロニー形成能で表すこともできる。すなわち、以下のような計算式に従って算出することができる。

10 コロニー形成能 = 導入したレプリコンRNAのコピー数 [コピー]/コロニー形成数 [個]

5. 本発明の他の実施形態

5

本発明に係るレプリコンRNA複製細胞は、例えばC型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングするための試験系として使用することもできる。具体的には例えば、被験物質の存在下で、レプリコン複製細胞を培養し、得られる培養物中のレプリコンRNAの複製を検出し、その被験物質がレプリコンRNAの複製を促進又は抑制するかどうかを判定することにより、C型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニングすることができる。この場合、得られる培養物中のレプリコンRNAの複製の検出は、レプリコンRNA複製細胞から抽出したRNA中のレプリコンRNAの量又は有無を検出することによるものであってもよいし、培養物中または該培養物に含まれるレプリコンRNA複製細胞中のタンパク質に含まれるHCVタンパク質の量又は有無を検出するものであってもよい。

このような本発明に係るレプリコンRNA複製細胞を用いる試験細胞系は、C型 25 肝炎ウイルス感染の治療のための治療剤若しくは診断剤の製造又は評価を目的と することが考えられる。そのような目的としては、具体的には、以下のような例 が挙げられる。

(1) 遺伝子型2aのHCVの増殖を抑制する物質の探索 遺伝子型2aのHCVの増殖を抑制する物質としては、例えば、直接的若しくは間

接的に遺伝子型2aのHCVの増殖に影響を及ぼす有機化合物、あるいは遺伝子型2aのHCVゲノム若しくはその相補鎖の標的配列にハイブリダイズすることによりHCVの増殖若しくはHCVタンパク質の翻訳に直接的又は間接的に影響を及ぼすアンチセンスオリゴヌクレオチド等が挙げられる。

5 (2) 細胞培養中で抗ウイルス作用を有する各種物質の評価

前記各種物質としては、合理的ドラッグデザイン又はハイスループットスクリ . ーニングを用いて得られた物質(例えば単離精製された酵素)等が挙げられる。

- (3) 遺伝子型2aのHCVに感染した患者の治療のための、新規攻撃標的の同定 例えばHCVウイルス増殖のために重要な役割を果たす宿主細胞性タンパク質を 同定するために、本発明に係るレプリコン複製細胞を使用することができる。
- (4) HCVウイルスの薬剤等に対する耐性獲得能の評価及び該耐性に関わる変異の 同定
- (5) C型肝炎ウイルス感染の診断薬又は治療薬の開発、製造及び評価のために使用可能な抗原としてのウイルスタンパク質の製造
- 15 (6) C型肝炎ウイルス感染の診断薬又は治療薬の開発、製造及び評価のために使用しうるHCVウイルス又はウイルス様粒子を製造するための、ウイルスゲノム複製系
 - (7) 遺伝子型2aのHCVに対するワクチンとして使用可能なワクチン抗原の製造
 - (8) 遺伝子治療用の外来遺伝子を組み込んで使用する、肝細胞指向性遺伝子ベク
- 20 ターの製造

25

10

6. 実施例

本発明を、以下の実施例及び図面に基づいてさらに具体的に説明する。但し、 本発明の技術的範囲はこれら実施例に限定されるものではない。

実施例1: レプリコンRNAの作製

(A) 発現ベクターの構築

劇症肝炎の患者から分離したC型肝炎ウイルスJFH-1株(遺伝子型2a)のウイルスゲノム全領域に対応するDNAを、該ウイルス株のゲノム全長cDNAを含むJFH-1

クローンから取得して、pUC19プラスミドに挿入したT7 RNAプロモーター配列の下流に挿入した。このようにして構築されたプラスミドDNAを、以下、pJFH1と称する。同様に、慢性肝炎の患者から分離したC型肝炎ウイルスJCH-1株(遺伝子型2a)のウイルスゲノム全領域に対応するDNAを、該ウイルス株のゲノム全長cDNAを含むJCH-1クローンから取得して、pUC19プラスミドに挿入したT7 RNAプロモーター配列の下流に挿入した。このようにして構築されたプラスミドDNAを、以下、pJCH1と称する。なお、上記JFH-1クローン及びJCH-1クローンの作製については、特開2002-171978号公報及びKato et al., J. Med. Virol., (2001) 64(3) p. 334-339に記載されている。またJFH-1クローンの全長cDNAの塩基配列は、国際DNAデータバンク(DDBJ/EMBL/GenBank)のアクセッション番号:AB047639に、JCH-1クローンの全長cDNAの塩基配列はアクセッション番号:AB047640に登録されている。

5

10

15

20

25

このようにして構築したプラスミドDNA pJFH1及びpJCH1の構造を、図1の上段に示す。「T7」はT7 RNAプロモーター、「G」は、挿入したJFH-1又はJCH-1由来DNAの5'端の上流かつT7 RNAプロモーター配列の3'端の下流に挿入したdGTP、を示す。「5'NTR」から「3'NTR」までがC型肝炎ウイルスのゲノム全領域に対応するDNAである。

次に、プラスミドDNA pJFH1及びpJCH1の構造領域と非構造領域の一部を、ネオマイシン耐性遺伝子(neo;ネオマイシンホスホトランスフェラーゼ遺伝子とも称する)及びEMCV-IRES(脳心筋炎ウイルスの内部リボゾーム結合部位)で置換して、プラスミドDNA pSGREP-JFH1及びpSGREP-JCH1をそれぞれ構築した(図1の下段)。この構築手順は、既報(Lohmann et al., Science, (1999) 285, p. 110-113)に従った。具体的には、プラスミドpJFH1及びpJCH1を制限酵素AgeI及びClaIで切断し、その切断部位に、pJFH-1由来の5'NTRからCore領域におよぶ配列とpRSV5NEO由来のネオマイシン耐性遺伝子とをPCR増幅により結合し制限酵素AgeIとPmeIで切断した断片、及びEMCV IRESからNS3領域におよぶ配列をPCR増幅により結合し制限酵素PmeIとClaIで切断した断片を、挿入し連結した。

また、pSGREP-JFH1中のNS5B領域について、該領域にコードされるRNAポリメラーゼの活性中心に相当するアミノ酸モチーフGDDをGNDに変異させる突然変

異を導入して、突然変異プラスミドクローンpSGREP-JFH1/GNDを作製した。

さらに、pSGREP-JFH1中のNS5B領域について、該領域にコードされるRNAポリメラーゼの活性中心に相当するアミノ酸モチーフGDDを含む連続した10アミノ酸配列を欠失させる突然変異を導入して、突然変異プラスミドクローンpSGREP-JFH1/dGDDを作製した。

なお上記で作製した突然変異クローンpSGREP-JFH1/GND及びpSGREP-JFH1/dGDD は、それらにコードされているNS5Bタンパク質の活性部位のアミノ酸配列が変異しているため、レプリコンRNAを複製するのに必要な活性NS5Bタンパク質を発現することができない。

10

15

20

25

5

(B) レプリコンRNAの作製

レプリコンRNA合成に用いる鋳型DNAを作製するために、上記のとおり構築した 発現ベクターpSGREP-JFH1、pSGREP-JCH1、pSGREP-JFH1/GND、pSGREP-JFH1/dGDD を、それぞれ制限酵素XbaIで切断した。

次いで、これらのXbaI切断断片のそれぞれについて、10~20μgを50μlの反応液中に含有させ、Mung Bean Nuclease 20 Uを用いて30℃で30分間インキュベートすることにより、さらに処理した。Mung Bean Nucleaseは、二本鎖DNA中の一本鎖部分を選択的に分解する反応を触媒する酵素である。通常、上記XbaI切断断片をそのまま鋳型として用いてRNA合成を行うと、XbaIの認識配列の一部であるCUGAの4塩基が3′末端に余分に付加されたレプリコンRNAが合成されてしまう。そこで本実施例では、XbaI切断断片をMung Bean Nucleaseで処理することにより、XbaI切断断片からCUGAの4塩基を除去した。この後、XbaI切断断片を含むMung Bean Nuclease処理後の溶液について、通常法に従ったタンパク質除去処理により、CUGAの4塩基が除去されたXbaI切断断片を精製して、これを鋳型DNAとした。次に、この鋳型DNAから、T7 RNAポリメラーゼを用いてRNAをin vitro合成した。このRNA合成にはAmbion社のMEGAscriptを用いた。鋳型DNAを0.5~1.0マイクログラム含む反応液20μlを製造業者の使用説明書に従って反応させた。

RNA合成終了後、反応溶液にDNase(2U)を添加して37℃で15分間反応させた後、さらに酸性フェノールによるRNA抽出を行って、鋳型DNAを除去した。この

ようにしてpSGREP-JFH1、pSGREP-JCH1、pSGREP-JFH1/GND、pSGREP-JFH1/dGDDに由来する上述の鋳型DNAから合成したRNA(レプリコンRNA)を、それぞれrSGREP-JFH1、rSGREP-JCH1、rSGREP-JFH1/GND、rSGREP-JFH1/dGDDと命名した。これらのレプリコンRNAの塩基配列を、rSGREP-JFH1については配列番号1及び図2A~F、rSGREP-JCH1については配列番号2及び図3A~F、rSGREP-JFH1/GNDについては配列番号7、rSGREP-JFH1/dGDDについては配列番号8に示す。

実施例2: レプリコン複製細胞クローンの樹立

5

10

15

20

(C) レプリコンRNAのトランスフェクション、トランスフェクション細胞のコロニー形成能の測定、及び細胞クローンの樹立

上記の合成レプリコンRNA(rSGREP-JFH1、rSGREP-JCH1、rSGREP-JFH1/GND、rSGREP-JFH1/dGDD)のそれぞれについて、様々な量のレプリコンRNAをHuh7細胞から抽出したトータル細胞性RNAと混合して、RNA総量が 10μ gとなるように調製した。次いでその混合RNAをエレクトロポレーション法によりHuh7細胞に導入した。エレクトロポレーション処理を行ったHuh7細胞を培養ディッシュに播種し、16時間から24時間培養した後に、培養ディッシュにG418(ネオマイシン)を様々な濃度で添加した。その後、週に2回培養液を交換しながら培養を継続した。播種時から21日間培養した後、クリスタルバイオレットで生存細胞を染色した。染色されたコロニー数を計測し、トランスフェクションしたレプリコンRNA量 1μ g当たりに得られたコロニー数を計算した。

コロニー形成が認められたrSGREP-JFH1又はrSGREP-JCH1トランスフェクション 細胞については、上記の培養21日後の培養ディッシュからさらに生存細胞のコロニーをクローン化し、培養を継続した。このようなコロニーのクローニングにより、細胞クローンを複数株樹立することができた。

25 樹立した細胞クローンについては、後述の実施例4と同様にして、複製レプリコンRNAの検出、ネオマイシン耐性遺伝子の宿主ゲノムDNAへの組み込みの有無の確認、及びHCVタンパク質の発現の確認を行った。細胞中でのレプリコンの複製が確認された細胞クローンを、レプリコン複製細胞クローンとした。

(D) 各トランスフェクション細胞におけるコロニー形成能

上記のトランスフェクションの結果、トランスフェクションしたレプリコン RNA $1 \mu g$ 当たりのコロニー形成能は、rSGREP-JFH1をトランスフェクションした Huh7細胞では、G418濃度が1.0 mg/m1の場合、94700 CFU(Colony Forming Unit; コロニー形成単位)/ μg ・RNAであった(図4の左列)。これに対して、rSGREP-JFH1/dGDD、rSGREP-JFH1/GNDをそれぞれトランスフェクションしたHuh7細胞では、コロニー形成が認められなかった(図4の中央列及び右列)。このことは、rSGREP-JFH1レプリコンRNAをトランスフェクションしたHuh7細胞のコロニー形成能は、rSGREP-JFH1から発現されるNS5B(RNAポリメラーゼ)の活性に依存していることを示す。つまり、コロニーを形成した細胞では、rSGREP-JFH1から発現されるNS5BのはたらきによりrSGREP-JFH1レプリコンRNAが自律複製することによって、ネオマイシン耐性遺伝子が持続的に発現されG418耐性が維持される結果、細胞増殖が可能になったものと考えられた。

一方、rSGREP-JCH1をトランスフェクションしたHuh7細胞では、G418濃度が $1\sim 0.5 mg/m1$ の場合にはコロニー形成が認められなかった(図5)。G418濃度を 0.25 mg/m1に下げた場合には、rSGREP-JCH1をトランスフェクションしたHuh7細胞でもコロニー形成が認められた。

さらに、上記(B)で得られた発現ベクターpSGREP-JFH1のXbaI切断断片をMung Bean Nucleaseで処理せずにRNA合成の鋳型DNAとして用いてレプリコンRNAを合成し、それを上記(C)と同様にしてHuh7細胞にトランスフェクションした。Mung Bean Nuclease処理を行わずに作製したこのレプリコンRNAには、CUGAの4塩基が3、末端に余分に付加されていた。

この結果、Mung Bean Nucleaseで処理せずに作製したレプリコンRNAをトランスフェクションしたHuh7細胞のコロニー形成能は、 $512~CFU/\mu~g$ ・RNAに低下した(図6の左側)。このことから、レプリコンRNAの3、末端の配列がトランスフェクションした細胞のコロニー形成能に影響を及ぼすことが明らかになった。

実施例3

5

10

15

20

25

(E) レプリコン複製細胞由来の複製レプリコンRNAの再トランスフェクション

実施例2に従ってrSGREP-JFH1のHuh7細胞へのトランスフェクションにより樹立したレプリコン複製細胞クローンから、常法により全RNA(トータルRNA)を抽出した。この細胞性RNAに含まれる複製レプリコンRNAのコピー数を、ノーザンブロット解析及び定量的RT-PCR法により決定した。

5

10

15

20

25

ノーザンブロット解析は、Molecular Cloning、A laboratory Manual、2nd edition、J. Sambrook、E.F. Fritsch、T. Maniatis著、Cold Spring Harbor Laboratory Press(1989)の記載に従って行った。細胞から抽出したRNAを変性 アガロース電気泳動に供し、泳動終了後に該RNAをポジティブチャージナイロン 膜に転写した。pSGREP-JFH1から作製した32PラベルしたDNAまたはRNAプローブを、前記のとおり膜に転写したRNAに対しハイブリダイゼーションさせ、次いでその 膜を洗浄し、それをフィルムに感光させることにより、レプリコン特異的なRNA バンドを検出した。

レプリコンRNAの定量的RT-PCRによる検出は、Takeuchi T, Katsume A, Tanaka T, Abe A, Inoue K, Tsukiyama-Kohara K, Kawaguchi R, Tanaka S, Kohara M. Real-Time detection system for quantification of Hepatitis C virus genome. Gastroenterology 116: 636-642 (1999) に従いHCV RNAの5'非翻訳領域のRNAを検出することによりおこなった。具体的には、細胞から抽出したRNAに含まれるレプリコンRNAを、合成プライマー、R6-130-S17、5'-CGGGAGAGCCATAGTGG-3'(配列番号13)、R6-290-R19、5'-AGTACCACAAGGCCTTTCG-3'(配列番号14)、TaqMan Probe、R6-148-S21FT、5'-CTGCGGAACCGGTGAGTACAC-3'(配列番号15)とEZ rTth RNA PCR kitを用いてPCR増幅し、次いでABI Prism 7700 sequence detector systemにより検出した。

次に、上記レプリコン複製細胞クローンのうちのクローン6、及びプールクローン (コロニー形成したレプリコン複製細胞を1ディッシュ分集めて培養した細胞) から抽出したトータル細胞性RNAの一部を、再トランスフェクションにより新たなHuh7細胞へ導入した。トランスフェクションに用いたトータル細胞性RNAは、上記で測定したレプリコンRNAのコピー数に基づき、1×10⁷コピーのレプリコンRNAを含むように調製した。トランスフェクションを上記(C)と同様の手法で行い、次いでG418濃度が1mg/m1の条件下で選択培養したところ、レプリコン

複製細胞のコロニー形成がみとめられた(図7)。この場合のコロニー形成能は、得られたコロニー数から計算すると、トランスフェクションに用いたレプリコン RNAの1×10⁶コピー当たり1コロニー以上であった。

一方、pSGREP-JFH1を鋳型としてT7 RNAポリメラーゼにより試験管内で合成したin vitro合成RNAのコピー数は、RNAの重量とRNAの長さから計算すると約 2×10^{11} コピー/ μ g・RNAであり、このin vitro合成RNAを上記と同様にしてトランスフェクションに用いた場合のコロニー形成能は、 5×10^{7} コピー当たり 1 コロニーであった。このような結果から、レプリコン複製細胞から抽出した細胞由来RNAとin vitro合成RNAを同じコピー数のレプリコンRNAとしてHuh7細胞にトランスフェクションした場合、Huh7細胞内で複製されたレプリコンRNAを用いると、コロニー形成能がin vitro合成RNAと比べて50倍程度高いことが明らかになった。

実施例4

5

10

15

20

25

(F) レプリコンRNAの検出

上記(E)に従ってrSGREP-JFH1のHuh7細胞へのトランスフェクションにより樹立したレプリコン複製細胞クローンから取得したトータルRNAを新たなHuh7細胞に再トランスフェクションして樹立した細胞クローン[クローン1~11]及びそれらのプールクローン (コロニー形成した細胞クローンを1ディッシュ分集めて培養した細胞)から、酸性フェノール抽出法によりトータルRNAを抽出した。次いでこのトータルRNAをノーザンブロット法により解析した。プローブとしてはpSGREP-JFH1特異的プローブを用いた。対照としては、トランスフェクションを行っていないHuh7細胞から同様に抽出したトータルRNA(図8中、「Huh7」として示す)、Huh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の7乗コピー加えたサンプル(図8中、「10⁷」として示す)、及びHuh7細胞から抽出したトータルRNAに試験管内で合成したレプリコンRNAを10の8乗コピー加えたサンプル(図8中、「10⁸」として示す)を用いた。図8中、1~11は細胞クローンの番号である。

この結果、rSGREP-JFH1と同程度の大きさのRNAがpSGREP-JFH1特異的プローブにより検出された(図8)。これにより、最初にトランスフェクションした

rSGREP-JFH1に由来するレプリコンRNAが細胞クローン内で複製増殖していることが確認された。また細胞クローン間で、複製レプリコンRNAの量に差があることが示された。図8中、例えば、クローン2、6,9、10は複製レプリコンRNAの量が多く、クローン4、8,11は複製レプリコンRNAの量が少なかった。

5

10

15

20

25

(G) ネオマイシン耐性遺伝子のゲノムDNAへの組み込みの有無の確認

実施例3の手順と同様にして作製した、レプリコンRNAを再トランスフェクシ ョンして得られた細胞クローンについて、その細胞クローンのG418に対する耐性 がネオマイシン耐性遺伝子のゲノムへの組み込みによるものでないことを確認す るために、ネオマイシン耐性遺伝子特異的プライマー(センスプライマー、NEO-S3:5'-AACAAGATGGATTGCACGCA-3'(配列番号16), アンチセンスプライマー、 NEO-R: 5'-CGTCAAGAAGGCGATAGAAG-3'(配列番号17))を用いて、細胞クローン から抽出した宿主細胞のゲノムDNAを鋳型とするPCR増幅を行った。用いた細胞ク ローンは、rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして 得られた細胞クローン1~8 (rSGREP-JFH1由来細胞クローン1~8)、及び rSGREP-JCH1由来の複製レプリコンRNAを再トランスフェクションして得られた細 胞クローン $1\sim6$ (rSGREP-JCH1由来細胞クローン $1\sim6$) であった。この結果、 図9に示すとおり、rSGREP-JFH1由来細胞クローンについては、調べた8クロー ン中、ネオマイシン耐性遺伝子の増幅が示された陽性クローンは認められなかっ た。rSGREP-JCH1由来細胞クローンについては、調べた6クローン中1クローン のみが陽性であった(図9中、右側写真のレーン3)。この陽性クローンは、 rSGREP-JCH1由来の複製レプリコンRNA中のネオマイシン耐性遺伝子が宿主細胞の ゲノムDNA中に組み込まれたことにより、G418耐性を獲得したと考えられた。こ の陽性クローンにおいては、他のクローンとは異なり、レプリコンRNA自体は細 胞内で自律複製していないものと考えられた。このことは、次の(H)に示す実験 で、この陽性クローンからHCVタンパク質が検出されなかったことにより確認さ れた。

(H) HCVタンパク質の検出

rSGREP-JFH1及びrSGREP-JCH1トランスフェクション細胞クローンから常法によ りタンパク質を抽出して、SDS-PAGE及びウエスタンブロット法により解析した (図10)。調べた細胞クローンは、上記(G)で用いたものと同じであり、 rSGREP-JFH1由来細胞クローン1~8、及びrSGREP-JCH1由来細胞クローン1~6 である。また、NS3遺伝子を含む発現プラスミドDNAをHuh7細胞にトランジエント 5 にトランスフェクションして得られた細胞抽出液を陽性対照(NS3タンパク質) とした。さらに、トランスフェクションしていないHuh7細胞から抽出したタンパ ク質を陰性対照として用いた。それぞれの細胞クローンから抽出したタンパク質 試料をPVDF膜 (Immobilon-P, Millipore社製) にブロッティングし、抗NS3特 異的抗体 (Dr. Moradpour より分与されたもの; Wolk B, et al, J. Virology. 10 2000; 74: 2293-2304) を用いて複製レプリコンRNAにコードされているNS3タン パク質を検出した。図10に示されるとおり、rSGREP-JFH1由来細胞クローン1 ~8及びrSGREP-JCH1由来細胞クローン1、2、4~6では、陽性対照と同じ大 きさのタンパク質が検出された。なおrSGREP-JCH1由来細胞クローン3 (上記(G)で陽性クローンとして検出されたクローン)では、NS3タンパク質の発現は検出 15 されなかった。すなわち、rSGREP-JCH1由来細胞クローン3では、レプリコンRNA の複製は確認されなかった。トランスフェクションしていないHuh7細胞でNS3タ ンパク質が検出されなかったため、NS3タンパク質が検出された細胞クローンで は、トランスフェクションされたレプリコンRNAが自律複製することによりNS3タ ンパク質が発現されていることが判明した。 20

なお、C型肝炎患者の血清を抗体として用いることにより、上記でNS3タンパク質の発現が確認された各細胞クローンについて、レプリコンRNAからのNS5aタンパク質の発現も確認した。

以上の(G)及び(H)の結果から、レプリコンRNAをトランスフェクションして 25 樹立した細胞クローンでは、レプリコンRNAが複製されていることが確認された。

実施例5

(I) 適合変異の解析

実施例3に従って、rSGREP-JFH1のHuh7細胞へのトランスフェクションを経て

樹立したレプリコン複製細胞クローンから取得したトータルRNAを新たなHuh7細胞に再トランスフェクションして、21の細胞クローンを樹立した。これらの細胞クローンから、常法によりトータルRNAをそれぞれ抽出した。このトータルRNAを鋳型にして逆転写酵素Superscript II(Invitrogen社製)とプライマー9641R-IH (5'-GCACTCTCTGCAGTCATGCGGCTCACGGAC-3'(配列番号18))によりレプリコンRNAに対応するcDNAを合成した。逆転写反応によるcDNA合成のための反応液組成を以下に示す。

		液量(µ1)
10	5x 1st strand Buffer	4
	2mM dNTP	5
	0.1M DTT	1 .
	9651R-IH プライマー (100μM)	1
	DW(蒸留水)	6. 5
15	Sample RNA (2 mg/mL)	1.
	RNasin (Promega社製)(40U/μL)	0.5
	Superscript II RT (Invitrogen社製)	1
	合計量	20 μ1

5

20 cDNA合成反応としては、まず上記のRNasinとSuperscript II以外の試薬を混合して最初の反応液を調製し、それを90℃で3分間加熱した後、氷上で冷却した。その後、この反応液に RNasinとSuperscript IIを添加して42℃で1時間反応させた後、さらに70℃で15分間反応させた。

さらに、このようにして得られたcDNAについて、以下の手順により5組のプラ 25 イマーセットを用いるPCR増幅を行って、レプリコンRNAのほぼ全領域にわたる DNA増幅断片を得た。用いたプライマーセット及びその各々により増幅される領 域を下記の表1及び表2に示す。

表 1

増幅断片の名称	プライマーセット		増幅領域
	プライマー1	プライマー2	
A/	42S-IH	433R-neo	41 - 470
B/	C/S17ssp	4680R-IH	28 - 3026
C/	4534S-IH	7279R-IH	2880 - 5625
D/	7198S-IH	9367R-IH	5544 - 7713
E/	9247S-NF	9576R-NF	7597 - 7960

なお表1中、増幅領域はrSGREP-JFH1(配列番号1)において対応する塩基番号で示した。

5

表 2

プライマーの名称	塩基配列 (5'→3')	配列番号
42S-IH	CCCCTGTGAGGAACTACTGTCTTCACGC	配列番号19
C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	配列番号20
4534S-IH	CCACTCAAAGAAAAGTGTGACGAGCTCGC	配列番号21
7198S-IH	GGCTTGGGCACGGCCTGA	配列番号22
9247S-NF	GCGGTGAAGACCAAGCTCAAACTCACTCCA	配列番号23
433R-neo	AGAACCTGCGTGCAATCCATC	配列番号24
4680R-IH	CCCGTCATGAGGGCGTCGGTGGC	配列番号25
7279R-IH	ACCAGCAACGGTGGGCGGTTGGTAATC	配列番号26
9367R-RI	GGCACGCGACACGCTGTG	配列番号27
9576R-NF	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	配列番号28

このPCR反応における反応液組成は以下のとおりである。

		液量(μ1)
	プライマー1(10μM)	1.0
	プライマー2(10μM)	1.0
5	2.5mM dNTPs	5. 0
	10x LA Buffer	5.0
	MgCl ₂ (25mM)	5.0
	LA Taq(TAKARA) (5U/ μ 1)	0.3
	DW (蒸留水)	30.7
10	铸型 cDNA	2.0
	合計量	50 μ1

また、PCR反応の条件は、以下のとおりであった: 95℃で2分間; 98℃で10 秒間に続き68℃で8分間を35サイクル; 72℃で7分間; 4℃で保持。

15 以上のようにして得られた各PCR産物の塩基配列を決定し、そのDNA配列対応するRNA配列とrSGREP-JFH1の配列との比較を行った。その結果を表3に示す。

表 3

領域	同義的 置換	非同義 置換	全変異
NS3	0	5	5
NS4A	0	2	2
NS4B	0	3	3
NS5A	0	7	7
NS5B	3	5	8
合計	3	22	25

20 表 3 のとおり、2 1 の細胞クローンにおいて認められた塩基変異は全部で2 5 個あったが、そのうち2 2 個はアミノ酸の変異を引き起こす非同義置換であった。

これらの変異の種類を、表 4 に示す。またこれらの変異の非構造領域における位置を図 1 1 に示す。

表 4

クローン名		変異部位		
	塩基番号	塩基変異	アミノ酸変異	アミノ酸番号
C1	7098	$A \Rightarrow G$	なし	
	7157	$A \Rightarrow G$	$Y \Rightarrow C$	2824
C2	4955	$c \Rightarrow 0$	$A \Rightarrow V$	2090
C3	4936	$A \Rightarrow G$	$T \Rightarrow A$	2084
	5000	$A \Rightarrow G$	$A \Rightarrow C$	2105
	7287	$A \Rightarrow G$	なし	
	7288	$A \Rightarrow G$	$M \Rightarrow V$	2868
C4	5901	$G \Rightarrow U$	$E \Rightarrow D$	2405
	6113	$A \Rightarrow U$	$H \Rightarrow L$	2476
C5	2890	$A \Rightarrow G$	$K \Rightarrow E$	1402
C6	7209	$A \Rightarrow G$. なし	

20 表4及び図11において、「C1~C6」は、変異が見出されたレプリコン RNAを有するレプリコン複製細胞クローンC1~C6を示す。「塩基番号」はレプリコンRNA rSGREP-JFH1の塩基配列(配列番号1)中の対応する塩基の番号として示している。「アミノ酸番号」は、JFH-1クローンにコードされるアミノ酸配列(配列番号4)中の対応するアミノ酸の番号として示している。変異部位の塩基及びアミノ酸の種類は、通常の表記法に従って記載されている。表4に示されるとおり、クローンC2では、レプリコンRNA上の配列番号1の塩基番号4955に相当する塩基がC(シトシン)からU(ウラシル)へ変異し、その結果、配列番号4のアミノ酸番号2090に相当するアミノ酸がA(アラニン)からV(バリン)へ変異した。

また図11に示した変異の位置は、表4に示す塩基番号を付記した縦線で示している。太い縦線は非同義置換、細い縦線は同義置換を表す。

なお、アミノ酸の変異を引き起こす塩基変異を全く有しないクローンも2クローンあった。この2クローンについてノーザンブロット解析を行ったところ、これらの2クローンにおいては、レプリコンRNAの複製量が、アミノ酸変異を生じる塩基変異を有するレプリコンRNAを複製している細胞クローンよりも少ないことが示された。つまり、レプリコンRNAにおけるアミノ酸変異を引き起こす塩基変異は、Huh7細胞においてレプリコンRNAの複製効率を増大させる適合変異であると考えられた。

10

15

25

5

実施例 6

(I) Huh7細胞以外の細胞を用いたレプリコン複製細胞クローンの樹立

実施例1に記載した方法に従って、Huh7細胞以外の肝癌細胞及び非肝臓由来細胞に、rSGREP-JFH1をトランスフェクションし、培養ディッシュ中に播種して培養して、コロニー形成の観察とコロニー数の計測を行った。使用した細胞は、以下のとおりである。

- (1) HepG2細胞 (Huh7細胞と共に代表的な肝癌細胞である)
- (2) IMY-N9細胞 (Itoらにより樹立された。HepG2細胞とヒト初代培養肝細胞との融合細胞; Hepatology 2001; 34: 566-572)
- 20 (3) Hela細胞 (ヒト子宮頸癌由来細胞; Can Cer Res. 1952; 12: 264-265)
 - (4) 293細胞 (ヒト胎児腎由来細胞; Gen. Virol. 1977; 36: 59-72)

HepG2細胞、IMY-N9細胞、Hela細胞、又は293細胞を用いたトランスフェクションの結果を、それぞれ図12A~Dに示す。図12A~Dに示すとおり、HepG2細胞、IMY-N9細胞、Hela細胞、及び293細胞はともに、rSGREP-JFH1トランスフェクション細胞におけるコロニー形成を示した。

樹立した細胞クローンは、後述の(L)、(M)と同様にして、複製レプリコン RNAの検出、ネオマイシン耐性遺伝子の宿主ゲノムDNAへの組み込みの有無の確認、

及びHCVタンパク質の発現の確認を行った。こうして細胞中でのレプリコンの複製が確認された細胞クローンを、レプリコン複製細胞クローンとした。すなわち、rSGREP-JFH1を用いることにより、今まで、HCVレプリコン複製細胞の作出に成功していなかった(Blight et al., Science, (2000) 290, 1972–1974)、Huh7以外の肝癌細胞、及び非肝細胞でのHCVレプリコン複製細胞の作製が可能であることが示された。

5

10

15

20

25

(K) Huh7細胞以外の細胞を用いたレプリコン複製細胞におけるレプリコンRNAの 検出

・ノーザンプロット解析は、 (Molecular Cloning, A laboratory Manual, 2nd edition, J. Sambrook, E.F. Fritsch, T. Maniatis 著、Cold Spring Harbor Laboratory Press (1989)) の記載に従って行った。前項(J)に従い、rSGREP-JFH1をHepG2、IMY-N9あるいはHeLa細胞へのトランスフェクションにより樹立し たレプリコン複製細胞クローン、及びrSGREP-JFH1を239細胞へのトランスフェク ションにより樹立したレプリコン複製細胞のプールクローン(コロニー形成した 細胞クローンを1ディッシュ分集めて培養した細胞)から、酸性フェノール抽出 法によりトータルRNAを抽出した。次いでこのトータルRNAをノーザンプロット法 により解析した。プローブとしてはpSGREP-JFH1特異的プローブを用いた。対照 としては、トランスフェクションを行っていないHuh7細胞及び、HepG2細胞から 同様に抽出したトータルRNA (図13中、レーン1及び17)、Huh7細胞から抽出 したトータルRNAに試験管内で合成したレプリコンRNAを10の7乗コピーあるいは、 10の8乗コピー加えたもの(図13中、レーン2及び3)を用いた。この結果、 rSGREP-JFH1と同程度の大きさのRNAがpSGREP-JFH1特異的プローブにより検出さ れた (図13)。これにより、最初にトランスフェクションしたrSGREP-JFH1に 由来するレプリコンRNAが細胞クローン内で複製増殖していることが確認された。 また細胞の種類により、複製レプリコンRNAの量に差があり、IMY-N9細胞は極め て効率よくレプリコンRNA を複製することが明らかになった。また、クローン間 で複製レプリコンRNAの量に差があることが明らかになった。

(L) ネオマイシン耐性遺伝子のゲノムDNAへの組み込みの有無の確認

樹立したレプリコンRNA複製細胞クローンについて、その細胞クローンのG418に対する耐性がネオマイシン耐性遺伝子のゲノムへの組み込みによるものでないことを確認するために、ネオマイシン耐性遺伝子特異的プライマー(センスプライマー、NEO-S3:5'-AACAAGATGGATTGCACGCA-3'(配列番号29)、アンチセンスプライマー、NEO-R:5'-CGTCAAGAAGGCGATAGAAG-3'(配列番号30))を用いて、細胞クローンから抽出した宿主細胞のゲノムDNAを鋳型とするPCR増幅を行った。用いた細胞クローンは、HepG2細胞に rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1、5、7、8、9、10、11、12、及び13、ならびにIMY-N9細胞に rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン3、4、5、6、7、8、9、10、及び11であった。この結果、図14に示すとおり、HepG2細胞にrSGREP-JFH1導入した細胞クローンについては、調べた9クローン中、ネオマイシン耐性遺伝子の増幅が示された陽性クローンは認められなかった。IMY-N9細胞にrSGREP-JFH1導入した細胞クローンについては、調べた9クローン中、ネオマイシン耐性遺伝子の増幅が示された陽性クローンは認められなかった。

同様に、Hela細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン及び293細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローンについて、同様の検討を行ったところ、ネオマイシン耐性遺伝子の増幅が示された陽性クローンは認められなかった。

(M) HCVタンパク質の検出

5

10

15

20

超立した細胞クローンから常法によりタンパク質を抽出して、SDS-PAGE及びウェスタンプロット法により解析した(図15)。調べた細胞クローンは、前項で用いたものと同じであり、HepG2細胞に rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1、5、7、8、9、10、11、12、及び13、ならびにIMY-N9 細胞にrSGREP-JFH1由来の複製レプリコ

 ν RNAを再トランスフェクションして得られた細胞クローン3、4、5、6、7、8、9、10、及び11である。また、既報(Lehmann et.al., Science, (1999))に従って、 μ HuH7にrSGREP-JFH1を導入して作製した μ HcV RNAレプリコン複製細胞クローンを陽性対照(図15 μ Pc 4-1、C6)とした。さらに、トランスフェクションしていない細胞から抽出したタンパク質を陰性対照として用いた(図15 μ Pc N)。それぞれの細胞クローンから抽出したタンパク質試料をPVDF膜(Immobilon-P, μ Millipore社製)にブロッティングし、抗NS3特異的抗体(Dr. Moradpour より分与されたもの;Wolk B,et al,J. Virology. 2000;74:2293-2304)を用いて複製レプリコンRNAにコードされているNS3タンパク質を検出した。図15上段に示されるとおり、rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1、5、7、8、9、10、11、12、及び13、ならびにIMY-N9細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン3、4、5、6、7、8、9、10、及び11では、陽性対照と同じ大きさのタンパク質が検出された。

5

15

20

25

また、C型肝炎患者の血清を抗体として用いることにより、上記でNS3タンパク質の発現が確認された各細胞クローンについて、レプリコンRNAからのNS5aタンパク質の発現の確認をおこなった。この実験では、患者血清を変更して抗NS5a 抗体を用いたこと以外は、NS3タンパク質の発現と同様にして検討を行った。その結果、図15下段に示すように、rSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン1、5、7、8、9、10、11、12、及び13、ならびにIMY-N9細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン3、4、5、6、7、8、9、10、及び11では、陽性対照と同じ大きさのタンパク質が検出された。

同様に、Hela細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローン及び293細胞にrSGREP-JFH1由来の複製レプリコンRNAを再トランスフェクションして得られた細胞クローンについて、同様の検

計を行ったところ、NS3及びNS5aタンパク質の発現が確認できた。

以上のとおり、レプリコンRNAをトランスフェクションして樹立した細胞クローンでは、レプリコンRNAが複製されていることが確認された。

5 実施例 7

10

15

(N) 適合変異の解析

実施例3に従って、rSGREP-JFH1のHepG2及びHeLa細胞へのトランスフェクションを経て樹立したレプリコン複製細胞クローンから取得したトータルRNAを新たな各々の細胞に再トランスフェクションすることにより、HepG2細胞については14の細胞クローンを、HeLa細胞については8の細胞クローンを樹立した。これらの細胞クローンから、常法によりトータルRNAをそれぞれ抽出した。このトータルRNAを鋳型にして逆転写酵素Supers Cript II(Invitrogen社製)とプライマー9641R-IH (5'-GCACTCTCTGCAGTCATGCGGCTCACGGAC-3'(配列番号31)によりレプリコンRNAに対応するcDNAを合成した。逆転写反応によるcDNA合成のための反応液組成を以下に示す。

	反応液組成	液量(μ1)
	5x 1st strand Buffer	4
	2mM dNTP	5
20	0.1M DTT	1
	9651R-IH プライマー (100μM)	1
	DW(蒸留水)	6. 5
	Sample RNA (2 mg/mL)	1
	RNAsin (Promega社製)(40U/μL)	0.5
25	Superscript II RT (Invitrogen社製)	1
	合計量	20 μ1

cDNA合成反応としては、まず上記のRNAsinとSuperscript II以外の試薬を混合して最初の反応液を調製し、それを90℃で3分間加熱した後、氷上で冷却した。

その後、この反応液に RNAsinとSuperscript IIを添加して42℃で1時間反応させた後、さらに70℃で15分間反応させた。

さらに、このようにして得られたcDNAについて、以下の手順により5組のプラ イマーセットを用いるPCR増幅を行って、レプリコンRNAのほぼ全領域にわたる DNA増幅断片を得た。用いたプライマーセット及びその各々により増幅される領域を下記の表5及び表6に示す。

表 5

増幅断片の名称	プライ・	プライマーセット	
1	プライマー1	プライマー2	
Α	42S-IH	433R-neo	41-470
В	C/S17ssp	4680R-IH	28-3026
С	4534S-IH	7279R-IH	2280-5625
D	7198S-IH	9367R-IH	5544-7713
E	9247S-NF	9576R-NF	7597-7966

10

なお表中、増幅領域は、rSGREP-JFH1(配列番号1)において対応する塩基番号で示した。

15

20

表 6

	プライマーの名称	塩基配列 (5' to 3')	配列番号
	43S-IH	CCCCTGTGAGGAACTACTGTCTTCACGC	配列番号32
5	C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	配列番号33
	4534S-IH	CCACTCAAAGAAAAGTGTGACGAGCTCGC	配列番号34
	7198S-IH	GGCTTGGGCACGGCCTGA	配列番号35
. •	9247S-NF	GCGGTGAAGACCAAGCTCAAACTCACTCCA	配列番号36
	433R-neo	AGAACCTGCGTGCAATCCATC	配列番号37
	4680R-IH	CCCGTCATGAGGGCGTCGGTGGC	配列番号38
10	7279R-IH	ACCAGCAACGGTGGGCGGTTGGTAATC	配列番号39
	9367R-IH	GGAACGCGACACGCTGTG	配列番号40
	9576R-NF	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	配列番号41

15 このPCR反応における反応液組成は以下のとおりである。

	反応液組成	液量(μ1)
	プライマー1 (10 μ M)	1.0
	プライマー 2 (10 μ M)	1.0
	2.5mM dNTPs	5.0
20	10x LA Buffer	5.0
	Mg C 12 (25mM)	5.0
	LA Taq(TAKARA) (5U/ μ 1)	0.3
	DW(蒸留水)	30. 7
	鋳型 CDNA	2. 0
25	合計量	50 μ1

また、PCR反応の条件は、以下のとおりであった: 95℃で2分間; 98℃で10秒間に続き68℃で8分間を35サイクル; 72℃で7分間;4℃で保持。

以上のようにして得られた各PCR産物の塩基配列を決定し、そのDNA配列対応するRNA配列とrSGREP-JFH1の配列との比較を行った。その結果を表7及び8に示す。

表7: HepG2 細胞におけるJFH-1 レプリコンにおける適合変異の解析

クローン	変	星部位	変	三異
_	塩基番号 アミノ酸番号		塩基	アミノ酸
HeplH1 HeplH3	6826 6887	2714 2734	C⇒A C⇒A	Q⇒K T⇒N
HeplH5	6887		U⇒A	なし
HeplH8	6580 7159	2632 2825	U⇒A U⇒C	S⇒T Y⇒H
HeplH9	3342 3594	·	A⇒G C⇒A U⇒A	なし なし N⇒K
HeplH10	7230 5052 6943	2848 2753	U⇒C C⇒A	なし P⇒T
HeplH12	なし			
HeplH13	4302 5687 6110	2334 2475	C⇒U G⇒A A⇒G	なし G⇒D Y⇒C

5

10

表7のとおり、HepG2細胞については8の細胞クローンにおいて認められた塩基変異は全部で13個あったが、そのうち8個はアミノ酸の変異を引き起こす非同義置換であった。これらの変異の種類を、表8に示す。一方、HeLa細胞については3の細胞クローンにおいて認められた塩基変異は全部で7個あったが、そのうち5個はアミノ酸の変異を引き起こす非同義置換であった。これらの変異の種類を、表8に示す。

PCT/JP2003/015038 WO 2004/104198

表 8: HeLa 細胞におけるJFH-1 レプリコンにおける適合変異の解析

クローン	変異部位			· · · · · · · · · · · · · · · · · · ·	
	塩基番号	アミノ酸番号		塩基	アミノ酸
HeLaH1	なし				
HeLalH2	5550	2272		U⇒C	S⇒P
	6252	•		A⇒G	なし
• •	7182			U⇒C .	なし
	7217	2844		A⇒G	H⇒R
HeLalH5	3643	1653		A⇒G	M⇒V
	5851	2389		G⇒A	A⇒T
	5914	2410	.)	G⇒A	E⇒K

表7において、「HepIH No.」は、HepG2 細胞を用いてクローン化したレプリ コンRNAを有するレプリコン複製細胞クローン番号を示す。「塩基番号」はレプ リコンRNA rSGREP-JFH1の塩基配列(配列番号1)中の対応する塩基の番号とし て示している。「アミノ酸番号」は、JFH-1クローンにコードされるアミノ酸配 列(配列番号4)中の対応するアミノ酸の番号として示している。変異部位の塩 基及びアミノ酸の種類は、通常の表記法に従って記載されている。表7に示され るとおり、例えば、クローン HepIH1では、レプリコンRNA上の配列番号 の塩基 10 番号 6826 に相当する塩基がCからAへ変異し、その結果、配列番号のアミノ酸番 号2714に相当するアミノ酸がQからEへ変異したことを示している。同様に表8に おいて、「HeLaIH No.」は、HeLa細胞を用いて、クローン化したレプリコンRNA を有するレプリコン複製細胞クローン番号を示す。

5

なお、アミノ酸の変異を引き起こす塩基変異を全く有しないクローンについて 15 ノーザンブロット解析を行ったところ、レプリコンRNAの複製量は、アミノ酸変 異を生じる塩基変異を有するレプリコンRNAを複製している細胞クローンよりも 少ないことが示された。つまり、レプリコンRNAにおけるアミノ酸変異を引き起 こす塩基変異は、細胞においてレプリコンRNAの複製効率を増大させる適合変異

であると考えられた。

産業上の利用可能性

本発明に係るレプリコン複製細胞は、遺伝子型2aのHCV由来のRNA及びHCVタンパク質を持続的に産生させるための培養系として利用することができる。さらに本発明に係るレプリコン複製細胞は、HCVの複製及び/又はHCVタンパク質の翻訳に影響を及ぼす各種物質をスクリーニングするための試験系として有用である。

本明細書中で引用した全ての刊行物、特許及び特許出願はその全体を参照によ 10 り本明細書中に組み入れるものとする。

配列表フリーテキスト

配列番号1-人工配列の説明:レプリコン

配列番号2-人工配列の説明:レプリコン

15 配列番号7-人工配列の説明:レプリコン

配列番号8~12-人工配列の説明:合成RNA

配列番号13~41-人工配列の説明:合成DNA

請求の範囲

1. 遺伝子型2aのC型肝炎ウイルスのゲノムRNA上の、5'非翻訳領域と、NS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする塩基配列と、3'非翻訳領域とを少なくとも含む塩基配列からなる、レプリコンRNA。

- 2. 少なくとも1つの選択マーカー遺伝子又はリポーター遺伝子、及び少なくとも1つのIRES配列を含む、請求項1記載のレプリコンRNA。
- 3. 配列番号9又は10で示される塩基配列からなる5'非翻訳領域と、少なくとも1つの選択マーカー遺伝子若しくはリポーター遺伝子と、IRES配列と、遺伝子型2aのC型肝炎ウイルスのゲノムRNA上のNS3タンパク質、NS4Aタンパク質、NS4Bタンパク質、NS5Aタンパク質及びNS5Bタンパク質をコードする塩基配列と、配列番号11又は12で示される塩基配列からなる3'非翻訳領域とを含む塩基配列からなる、レプリコンRNA。
- 15 4. 遺伝子型2aのC型肝炎ウイルスのゲノムRNAが、配列番号3又は5で示される塩基配列からなるRNAである、請求項1~3のいずれか1項記載のレプリコンRNA。
 - 5. 以下の(a)又は(b)のRNAからなるレプリコンRNA。

5

- (a) 配列番号1又は2で示される塩基配列からなるRNA
- 20 (b) 配列番号1又は2で示される塩基配列において1~10個の塩基が欠失、 置換又は付加された塩基配列からなり、かつ、自律複製能を有するRNA
 - 6. 請求項1~5のいずれか1項記載のレプリコンRNAを細胞に導入すること により作製された、レプリコン複製細胞。
 - 7. 細胞が真核細胞である、請求項6記載のレプリコン複製細胞。
- 25 8. 真核細胞がヒト肝由来細胞、ヒト子宮頸由来細胞、又はヒト胎児腎由来細胞である、請求項7記載のレプリコン複製細胞。
 - 9. 真核細胞が、Huh7細胞、HepG2細胞、IMY-N9細胞、HeLa細胞、及び293細胞 からなる群より選ばれるいずれかの細胞である、請求項7記載のレプリコン複製細胞。

10. C型肝炎ウイルス感染の治療剤若しくは診断剤の製造、又は評価のための、 請求項 $1\sim5$ のいずれか1項記載のレプリコンRNA。

- 11. C型肝炎ウイルス感染の治療剤若しくは診断剤の製造、又は評価のための、 請求項6~9のいずれか1項記載のレプリコン複製細胞。
- 5 12. C型肝炎ウイルス感染に対するワクチンの製造のための、請求項1~5の いずれか1項記載のレプリコンRNA。
 - 13. C型肝炎ウイルス感染に対するワクチンの製造のための、請求項6~9の いずれか1項記載のレプリコン複製細胞。
- 14. 請求項6~9のいずれか1項記載のレプリコン複製細胞からレプリコン10 RNAを抽出することを含む、遺伝子型2aのC型肝炎ウイルスのレプリコンRNAの製造方法。
 - 15. 請求項6~9のいずれか1項記載のレプリコン複製細胞を培養し、得られる培養物からウイルスタンパク質を取得することを含む、遺伝子型2aのC型肝炎ウイルスのウイルスタンパク質の製造方法。
 - 15 16. 被験物質の存在下で、請求項6~9のいずれか1項記載のレプリコン複製 細胞を培養し、得られる培養物中のレプリコンRNAの複製を検出することを 含む、C型肝炎ウイルスの複製を促進又は抑制する物質をスクリーニング する方法。
 - 17. 請求項6~9のいずれか1項記載のレプリコン複製細胞から複製レプリコ ンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは 別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上 行うことを含む、遺伝子型2aのC型肝炎ウイルスのレプリコンRNAの複製効率を増大させる方法。
 - 18. 複製効率の増大が、レプリコン複製細胞に最初に導入されたレプリコン
 25 RNAの複製効率と比較して、少なくとも2倍の増大である、請求項17記載
 の方法。
 - 19. 請求項6~9のいずれか1項記載のレプリコン複製細胞から複製レプリコンRNAを取得し、取得した複製レプリコンRNAを該レプリコン複製細胞とは別の細胞に導入して新たなレプリコン複製細胞を作製する工程を1回以上

行うこと、及び最終的に得られたレプリコン複製細胞から複製レプリコン RNAを取得することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。

- 20. 請求項19に記載の方法によって製造された複製効率が増大したレプリコンRNAについて、レプリコン複製細胞に最初に導入されたレプリコンRNAとの間の塩基変異又はアミノ酸変異を検出すること、及び複製効率を増大させようとするレプリコンRNAにその検出された塩基変異又はアミノ酸変異を導入することを含む、複製効率が増大した遺伝子型2aのC型肝炎ウイルスのレプリコンRNAを製造する方法。
- 10 21. 配列番号1で示される塩基配列上において、以下の(a)~(u):
 - (a) 塩基番号7157の部位におけるAからGへの変異、
 - (b) 塩基番号4955の部位におけるCからUへの変異、
 - (c) 塩基番号4936の部位におけるAからGへの変異、
 - (d) 塩基番号5000の部位におけるAからGへの変異、
 - (e) 塩基番号7288の部位におけるAからGへの変異、

15

20

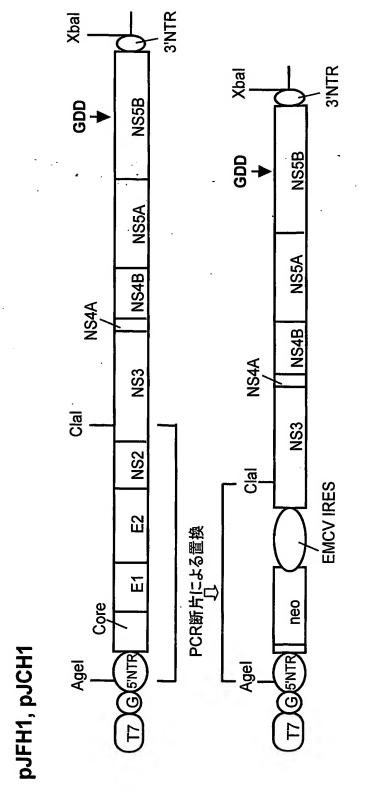
25

- (f) 塩基番号5901の部位におけるGからUへの変異、
- (g) 塩基番号6113の部位におけるAからUへの変異、
- (h) 塩基番号2890の部位におけるAからGへの変異、
- (i) 塩基番号6826の部位におけるCからAへの変異、
- (j) 塩基番号6887の部位におけるCからAへの変異、
- (k) 塩基番号6580の部位におけるUからAへの変異、
- (1) 塩基番号7159の部位におけるUからCへの変異、
- (m) 塩基番号7230の部位におけるUからAへの変異、
- (n) 塩基番号6943の部位におけるCからAへの変異、
- (o) 塩基番号 5 6 8 7 の部位におけるGからAへの変異、
 - (p) 塩基番号6110の部位におけるAからGへの変異、
 - (a) 塩基番号 5 5 5 0 の部位におけるUからCへの変異、
 - (r) 塩基番号7217の部位におけるAからGへの変異、
 - (s) 塩基番号3643の部位におけるAからGへの変異、

(t) 塩基番号5851の部位におけるGからAへの変異、及び

(u) 塩基番号 5 9 1 4 の部位におけるGからAへの変異、からなる群より選択される少なくとも1 つの変異を有する塩基配列からなるレプリコンRNA。

5



pSGREP-JFH1, pSGREP-JCH1

図2A

03 UEUDAUDAAD	CCCUGUGAG	40 AUGAAUCACU	30 ACACGCOGCC	20 AAUAGGGGCG	ACCUGCCCCU
120 CUCCAGGCCC	110 GUCGUACAGC	100 Daguadgagu	90 GCCAUGGCGU	08 ADDOGOGAGA	70 CUUCACGCAG
AAUUGCCEGG	170 AGUACACOSG	160 GGAACCEGOG	150 AGUGGUCUGC	140 GGAGAGCCAU	CCCCCUCCOG
ceuecoccos	230 GCCAUUUGEG	220 UCUADGCCCG	AUAAACCCAC	200 CCUUUCUUGG	190 AAGACUGGGU
300 CCUGAUAGGG	290 UGUGGUACUG	280 CGAAACGCCU	CGUUGGGUUG	260 AGCOGAGUAG	250 CAAGACUGCU
350 AUCCUAAACC	350 AUGAGCACAA	340 ACCGUECACC	330 GGUCUCGUAG	320 Geoggagga	OCCUUGOGAG
420	410	400		380	370
480	470	460	450 GGGUGGAGAG	440	430
. 540	530	520		500	890
600	:590	. 580	•	560	550
. 660	650	540	•	620	610
720	710	700	690 GCGAAGUGGC	. 580	670
	770 CGGCUGCAUA	760 UGCAAUGCSG	750 UCAUGGCUGA	740 AAAGUAUCCA	730 UCCUGCOGAG
	058 GAGCGAGCAC	ACAUCGCAUC	ACCAAGCGAA	600 CCAUUCGACC	790 GGCUACOUEC
		880 GGACGAAGAG	870 AGGAUGAUCU	860 CURIGUCGAUC	850 GGAAGCÖGGU
960 COCUÇAÇÕES	950 GAGGAUCUCG	940 GCCCGACGGC	930 AGGCGCGCAU	920 GCCAGGCDČA	910 CGAACUGUUC
L020 GAUECAUCGA	1010 CGCUUUUCUG	1000 GGAAAAUGGC	990 AUAUCAUGGU	089 AECCOUUCOUA	970 UGGCGAUGOD
1080 CCCGUGAUAU	1070 GCGUUGGCUA	1060 UCAGGACAUA	1050 CGGACCGCUA	1040 CUGGGUGUGG	1030 CUGUGGCCGG
1140 GUAUOGCOSC	GUGCUUUNCG 1130	1120 CCCCUUCCUC	1110 AAUGGGCUGA	1100 CUUGECGGOB	1090 UGCUGAAGAG
1200 GAGUUUAAAC	1190 GAGUUCUUCU	2180 CCUUCUUGAC	1170 CCUUCUAUCS	1160 CAGCGCARCG	1150 UCCCGAUUCG
Gecoenene 1360	1250 CUUGGAAUAA	1240 GCCGAAGCCG	AACGUUACUG	1220	1210
	1310	1300	1290	1280	1270
	1370	1360	1350	1340	1330
					· ~ · · · · · · · · · · · · · · · · · ·

図2B

•					
1398 DGCAAGGUCU	1400 Guugaauguc	1410 GUGAAGGAAG	1420 CAGUUCCUCU	1430 GGAAGCUUCU	1440 Ugaagacaaa
1450 CAACGUCUGU	1460 AGCGACCUU	1470 UGCAGGCAGC	1480 GGAACCCCCC	1490 ACCUGGCGAC	1500 AGGUSCCUCU
	4000	1530 UAAGAUACAC	4646	1550	1560
1570 UUGUGAGUUG	1580 GADAGUUGUG	1590 GAAAGAGUCA	1600 AAUGGCUCUC	1610 CUCAAGOGUA	1620 UÇCAACAAGG
	1.540	1650 GUACCCCAUU	1660	1670	1680
1590	. 3700	1710 UCGAGGUUAA	. 1720	1730	1740
1750 ACGUGGUUU	1760 CCUUUGAAAA	1770 ACACGAUGAU	1780 ACCAUGGCUC	1790 CÉAUCACUGO	1800 UUAUGCCCAG
1810 CAAACACGAG	1870 CCCUCCUGGG	1630 CGCCAUAGUG	1840 GUGAGUAUGA	1850 OGGGGCGUGA	1860 CAGGACAGAA
1870 CAGGCCGGG	1680 Aaguccaaau	1890 CCOGUCCACA	1900 GUCUCUCAGU	1910 CCUUCCUCGG	AACAACCAUC
UCCGGGGUUU	UGUGGACUGU	1950 UUACCACGGA	GCUGGCAACA	AGACUCUAGC	CEGCUUACIAG
GGUCCGGUCA	CGCAGAUGUA	CUCGAGUGCO	GAGGGGGACU	UGGUAGGCUA	GUCLIAGULAGU
2050 CCUGGGACCA	2060 AGUCUUUGGA	2070 GCCGUGCRAG	2090 UGUGGAGCOG	UCGACCUADA	UCUGGUCACG
CEGAACECUG	AUGUCAUCCC		CCCCCCCACA	AGCGGGGGG	AULGEOCUCU.
2170 CCGAGACCCA	DUUCGACCUU	GAAGGGGUCC	DOGGGGGGC	OGGUGCUCUG	CCCOMOSSEC
CFOGNOCALE 3530	2240 GCCUCUUCCG	2250 AGCAGCUGUG	vecucucese vecucucese	2270 GOGUGGOCAA	2280 AUCCAUCGAU
2290 EDCAUCCCCG	2300 UUGAGACACU	CGACGUUGUU	2320 ACAAGGUCUC	2330 CCACUUUCAG	2340 UGACAACAGC
	CUGUGCCCCA		GUCGGGUACU	UCCAUCCUCC	2400 AACIXGGCAGU
GBAAAGAGCA	CCAAGGUCCC		GCCGCCCGGG	General	MONINGARAN
					2520 ACAUGECAUC
2530 AAUCCCAÄCA	2540 UUAGGACUGG	2550 AGUCAGGACC	08ES GUGAUGACOG	2570 GGGAGGCCAU	2580 CACGUACUCC
2590 ACAUAUGGCA	2500 AAUUUCUOGO	2610 CGAUGGGGGC	3620 DCCCUAGCO	2630 GOGCCUAUGA	2640 CAUCAUCAUA
2550	2680	2670	2680	2690	2700 GGUÇCUUGAU
2716	2730	2730	2740	2750	2760
CAAGCAGAGI	CAGCCGGGG	CAGACUAACU	GUGCUGGCUA	CGGCCACACC	CCCCGGGUCA

図2C

2820	2810	GUAGGCCUCG	0e7S	2780	2770
UGAGAUCCCC	GGCGGEAGGG		Dadagagadau	CCCAUCOCGA	GUGACAACOC
) GAUUUUCUGC	2870	2860	0585	GGGCGADUCC	2830
	GGAGACACCU	AUCAAGGGAG	OCUAUCOUGO	2840	UUCUAUGGGA
2940	. 7930	2920	2910	2900	2890
CUUGAAUGCC	GEGGCAUGGG	GCGGCCCUUC	CGAGCUCGCG	Aaaguguga	CACUCAAAGA
Dedeedeede	2990	2980	2970	2960	S950
	CUCAGGGAGA	AUAAUACCAG	GGACGUCUCC	Aurorgood	COĞGCAUACU
3060	3050	3040	3030	3020	3010
GAUCGACUGC	UUGACUCĆGU	ACUGGAGACU	GACGGGGUAC	AGSCCCUCAU	GUCGCCACO3
3120	3110	3100	3090	3080	3070
UNUAACCACA	CCACCUUCAC	AGCCTGGACC	UGUCKACUUC	DCACCCAAGC	AMUGUAGOGG
3180	3170	3160	3150	3140	3130
AGGUAGAGGA	GOGGGGGGCAC	AGUCAGOGCC	UGUCUCACGC	CACAAGACGC	CAGACUGUCC
3240	3230	3220	3210	3200	3190
GUUGACAGU	OCUCAGGAAU	GGUGARCGAG	UGUUUCCACU	CULIAUAGGUA	AGACAGGGCA
3300	3290	3280	3270	3260	3250
ACCAGCGAG	ACGAUCUCAC	GCUGCGUGGU	CGACGCAGGG	GUGAGUGCUA	GUAGDECUUU
3360	3350	3340	3330	3320	3310
UCARGACCAU	UACCCGUGUG	AOGCCCCGGCC	GUADUUCAAC	GGCUUAGAGC	ACCACOGUCA
3420	J410	3400	3390	3380	2370
COUCCOCOCC	UAGAOGCCCA	CUCACACACA	UUUCACCGGC	GGGAGGCAGU	CUUGAAUUUU
3480	3470	3450	3450	0106	3430
DACGGUGUGC	CCUACCAAGC	UACCUAGUAG	GAACUUCGCG	Abgogggga	CAAACAAAGC
3540 GGCCCGACUC					
3600 UACCAADGAG	OGGGCCCUAU	CUGUACCGUU	CRCACCUCUC	Onedeleccc	AAGCCUACGC
3660	3650	3640	OE&E	3620	3610
UGACCUUGAG	GCAUGCAAGC	AUCGCCACAU	GACGAAGUAC	CACACCCUGE	GUCACCCUCA
3720 CGCAUAUUGC	CYCCCGACCC	GGAGUCCUGG	3690 ADDUDDAUDD	3680 GCACGUEGGU	3670 GUCAUGACCA
3780 AGUCGUCGUU	UCAACCAGCG	CECUUGCACG	CAUCAUCGGC	GAUGCGUUUC	3730 COGGOGACUG
3840 OGCCUCUAGG	UGGAGGAAUG	UUUGAUGAGA	GUAUGAGGCU	ACCAGGUCCU	
3900	0e8é	3880	3870	3860	3950
GAUCCAAGGC	Aaccucaagu	GCCGAGADGU	GCAGOGGAUA	UCGAAGAGGG	GCGGCUCUCA
3960	3950	3940	3930	3920	3910
GCCUUCAUGG	CCCCUAUGCA	GACAUACAAC	GCAGGCCCAG	AGGCCUCUAA	UUGCUGCAGC
020	4010	adda	3990	3980	3970
CAUCCAAUAC	UCAUUAGCGG	Avguegaacu	GGCCAGACAC	AACAAUUUUG	CCCAAAGUGG
4080	4070	4050	4050	4040	4030
GSCAUUCĀGU	CUUCCAUGAU	CCCGCGGGG	GCCAGGGAAC	UGUCAACACU	COCCAGGAU
4140	4130	ACCACCAUCC	4110	41.00	4090
CAUGGGAGGC	UUCUCAACAU		GUCGACCAGU	CCAGUCCGUU	GCCGCCCUCA

図2D

4200	4190	4180	ACCACCCGGG	4160	4150
CAGUGGCCUG	6CUIDGUCGU	GGGCCACCG		CCCAGAUCGC	UGGUUAGOGU
4250	4250	4240	4230	4220	4210
CCUGGCAGGA	UGGUGGAÇAU	Geuaagguge	CAUAGGCCUG	COGUGGGCAG	GUGGGGGCUG
ÇGAGAAGCCC	4310	4300	4290	4280	4270
	UCAUGUCUGG	GCAUUCAAGA	GGCCUCGUC	GCAUUUCGGG	UAUGGUGCGG
4380	4370	eggauccugu	4350	4340	4330
CCUGGUGGUG	CUCCGGGAGC		UCUACUGCCU	AUGUCAUCAA	UCUAUGGAAG
	4430	4420	4410	4400	2390
. 4500 UACUCACUAC	4490	4480	4470	4460	4450
4560	4550	4540	4530	4520	4510
UACUAUAACC	UUGGCUCUCU	ACCCAACUAC	GCAGCGUGUG	COGADOCOUC	GUGACEGAGU
4620 AUGCUCCGGA					
4480 Caaaaauugg					
4740 UCAAAAGGGG	4730	4720	4710	4700	4690
4800	4790	4780	4770	4760	4750
CGGCGCCAAC	GCUGCCCUUG	AUGACCACGC	CACUGGCAUC	UGUGGGCCGG	UACAAGGGUG
4860	4850	4840	4830	4820	4810
CUGCAUGAAC	GGCCUAAAAC	AGGAUCACAG	GGGCUCUAUG	AUGUCOGOCU	AUCUCUGGCA
GCCGAAACCC	4910	4900	4890	ARRO	4870
4980	4970	4960	4950	4940	4930
GGAGGUGACG	OGGAGUAOGC	GUGGCGGCCU	CAUCUGGAGG	ACAAGACCEC	CCCACGAACU
5040	5030	5020	5010	5000	d990
AAUUCCUUGC	ACAAUCUGAA	CUGACCACUG	UGUAACAGGA	OGUACUCCUA	CAGCAUGGGU
5100	5090	0805	5079	SOEO	5050
VAGGUUUGCA	UGCAGAUCCA	DUDGOADDUD	UUUCUCCUGG	CUCCAGAGUU	CAACUACCUU
5160	5150	5140	9130	5120	5110
VAAUUOCUAU	GCGUUGGGCU	GUCUCGUUCU	COGGGAUGAG	AGCCGUUUUU	CCCACACCAA
5220	5210	5200	5190	5180	5170
GAGGUCCAUG	CAGACGUAUU	GAGCCCCGACG	CUGUGAACCU	CCCAGCUUCC	GCUGUCGGGU
5280	5270	5260	<i><u>C</u>yÓ</i> 66CGCBG	5240	5330
ACGGGGAUCA	GCCCUUGGC	ACUGOGGGG	2320	OGCCCCAÇAU	CUAACAGAUC
5340	5330	5320	5310	AGECGAGCUC	5290
GOGGGCCACC	CACCGUCGCU	CAGCUAUCAG	CUCAGUGAGC		CCUCCAUCUG
5400	5390	5380	5370	5360	5350
GCUCAUGGAG	AUGCCAACCU	GACAUGGUCG	CUAUGACGUG	ACAGCAACAC	153CACCACCC
	5450	5440	5630	5420	5410
•	•			•	
6520CCCCAGG	CGGAGUGCAU	UCAAUACCAU	5490 CCUUGAĞCÇC	nere Adoranaaa Adoranaaaa	9470 DDADDDDUA

図2E

S530						
5550 5660 5670 5680 5690 5770 5710 5720 5730 5730 5750 5810 5820 5820 5820 5820 5820 5850 5860 5870 5860 5870 5860 5860 5870 5860 5860 5860 5860 5860 5860 5860 5960 5960 5960 5970 5980 6000 60	5580 GCCGCUCGUG	5576 ACUACAACCC	5560 GCACGGCCUG	9350 ACCGGCUUGG	5540 CACGGGCCUU	5530 AGOGGGUUUC
1710 5720 5730 5740 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5750 5760 5770 5770 5780 5790 5890 5810 5820 5830 5840 5850 5860 5870 5860 5870 5860 5890 5900 5910 5920 5930 5940 60000000000000000000000000000000000	5640 UCUCCCCCC	5630 CUGGUUGUGC	5620 CCCACCGUUG	5610 UUACCAACCG	S600 GGAGGCCAGA	5590 GAAUGGUGGA
\$770 \$780 \$790 \$800 \$810 \$820	GAGCGAGAGC	CAGUGGGUGU	AGACGCCGGA	UCCCCCAAGG	CCCCGACGCC	CCCAAGAAGG
\$770 \$780 \$790 \$800 \$810 \$820	5760 CCCCUCGAGC	5750 UUGGČCAGCC	5740 AUCAAGACCU	5730 GCAACUGGCC	5720 AAGCCCUCCA	5710 ACCAUAUCAG
1990 5990	5820	5810	5800	5790	5780	5770
Secondario	OBBE OGAÇOGOGA OGAÇOGOGA	5870 UGOCCCCCCU	5860 GCCUCCUCUA	5850 GACAGGUUCC	5840 CCCCUCAGA	5830 GGUGAGCOGG
6010 6020 6030 6040 6050 6050 6050 6050 6050 6050 605	CCYCCCCCCC	AACCUCCCCC	GUAGAGCUUC	GUCUGAUCAG	CGGACCUGGA	CCUGGAGAUC
CONTROL CONT	GGACGAUACC	GCUCCGAGGA	UGGUCUACUU	CUCGGGGGCC	CCCGUUCGGG	GEGGUAGCUC
AGAGGRAA AGUUGCCAAU CAACCCUUG AGUAACUCCC OGUUGCGAUA CCAUAACAAG 5130 6140 5150 6160 6170 UUUGACAGG 6190 6200 6210 6220 6230 6240 ACACAGGCCCA UUAUGACUCA GUCUUAAAGG ACAUCAAGCU AGCGCCUUCC 6250 6260 6270 6280 6290 6290 ACCUCAGUCC AGGUCAGCC CAAGGCUCCU CACCUUGGAG GAGGCCUGCC AGUUGACUCC ACCCCAUCUCC 6310 6320 6330 6340 6350 6360 6360 AAGAGUCCA AGUAUGGAUU CGGGCCCAAG GAGGUCGCA GCUUGCCCGG GAGGGCCGUUCC 6370 6380 6390 6400 6400 6410 6420 ACCACAUCA AGUUCCGUGGG GAGGACCCCG CCAAAACACC AAUUCCCACAC 6430 6440 6450 6460 6470 6480 ACAUCAUGG CCAAAAAUGA GGUGUUCUGC GUGGACCCCG CCAAGGGGC UAAGAAACCA 6430 6500 6510 6520 6530 6540 ACCCCCUCA UCCGUUCCCC UGACCUCGC GUCCGGGUCU GCGAGAAAAU GGCCCCUUAU 6550 6560 6570 6580 6590 6600 ACAUUACAC AAAACUCC UCAGCCGGUC GUCCGGCCU CCUAUGGCCU CCAGUACUCC 6610 6620 6630 6640 6550 6560 ACAUUACAC GGGUGGAGAU UCUCUUGAAA GCAUCGGCC AAAAAAACC CAGUACCCC 6670 6680 6690 6700 6710 6720 ACCCCAAAA ACCCGAUG CUUCCACUCA ACCGUCACUC AGAAAACAC CAGGACCGAG AUUCCCUAU ACCCCAAC GGGGCCCCG CCAAGAACAC CAGGACCGAG ACCCCAACAC GGGUGGACUC CUUCCACUCA ACCGUCACUC AGAAAACAC CAGGACCGAG ACCCCAACAC GGGUGGACUC CUUCCACUCA ACCGUCACUC AGAAAACAC CAGGACCGAG ACCCCAAC GGGUGGACUC CUUCCACUCA ACCGUCACUC ACAAACAC CAGGACCGAG ACCCCAAC GGGUGGACUC CUUCCACUCA ACCGUCACUC ACAAACAC CAGGACCGAG ACCCCAACAGACAC CUUCCACUCA ACCGUCACUC ACCGCCCAU ACCCCCAUGCCAU ACCCCAACAC GGGUGGACCC CUUCCCCCC GAGGACGCC CCACUGCCAU ACCCCCAUGCCAU ACCCCAACAC CUUCCACUCA ACCGUCACUC ACCGCCCAUGCCAU ACCCCCCAUGCCAU ACCCCCAUGCCAU	CUGUAGCCCC	DAAVAACUCC	ACCEGGGCUC	AUACUCCUGG	GCUCCAUGUC	ACCGUGUGCU
6190 6200 6210 6220 6230 6240 ECCAAGUCC UCACGCCCA UUADGACUCA GUCUUAARGG ACAUCAAGCU ACCCCAUUCC 6250 6260 6270 6280 6290 6390 EGUCAGCC CAAGGCUCCU CACCUUGGAG GAGGCCUCCC AGUUGACUCC ACCCCAUUCU 6310 6320 6330 6340 6350 6360 EAGAUCCA AGURUGGAU CCGCGCCAAG GAGGUCCGCA GCUUGUCCGG GAGGCCCUU CCACAUCA AGUCUGGAG GAAGGCCCC CUGGAAGACAC CACAAACACC AAUUCCCACA CCACAUCA AGUCCGUGUG GAAGGACCUC CUGGAAGACC CACAAACACC AAUUCCCACA CCACAUCA CCACAAACACA GGUUUCUGC GUGGACCCCC CCAAGGGGGG UAAGAACCA CCACAUCA CCACAAACACA GGUUUCUGC GUGGACCCCC CCAAGGGGG UAAGAACCA CCACAUCA CCAAAAAUGA GGUUUCUGC GUGGACCCCC CCAAGGGGG UAAGAACCA CCACAUCA UCCUUACCC UGACCUCGCC GUCCGGGUCU GCGAGAAAAU GGCCCCUCUAU CCCCCCCCA UCCGUUACCC UCAGGCGGGA ACCGCGCGG AAAAGAACAC CCACAGACCCC CCACGGGCUUCCC UCAGGCCGUA ACCGCCCGG AAAAGAAGAA CCCCCCCCCC	CCAUAACAAG	OGUUGCGAUA	AGUAACUOGC	CAACCCUUUG	AGUUGCCAAU	GAAGAGGAAA
### CAAGUCC UCGACGCCCA UUAUGACUCA GUCUUAAAGG ACAUCAAGCU ACCCCAUUCC 6250						
6310 6320 6330 6340 6350 6360 CARGAUCCA ACUAUGGAUU CGGGGCCAAG GAGGUCCGCA GCUUGUCCGG GAGGGCCGUU 6370 6380 6390 6400 6410 6420 CCACAUCA AGUCCGUGU GAAGGACCUC CUGGAAGACC CACAAACACC AAUUCCCACA 6430 6440 6450 6450 6460 6470 6480 CCACAAAAAUGA GGUGUUCUGC GUGGACCCCG CCAAGGGGGG UAAGAAACCA 6490 5500 6510 6520 6530 6540 EUGGCCUCA UCCUUUACCC UGACCUCGGC GUCCGGGUCU GCGAGAAAAU GGCCCUCUAU 6550 6560 6570 6580 6690 6690 6600 CCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	Veceeconoc	ACAUCAAGCU	GUCUUAAAGG	UUAUGACUCA	UCGACGCCCA	ACCCAAGUCC
6370 6380 6390 6400 6410 6420 CCACAUCA AGUCCGUGUG GAAGGACCUC CUGGAAGACC CACAAACACC AAUUCCCACA 6430 6440 6450 5460 6470 6480 CCACAUCA CCAAAAAUGA GGUGUUCUGC GUGGACCCCG CCAAGGGGGG UAAGAAACCA 6490 6590 6510 6520 6530 6530 6540 EUCCCUCA UCGUUUACCC UGACCUCGGC GUCCGGGUCU GCGAGAAAAU GGCCCCUCUAU 6550 6560 6570 6580 6590 6690 6690 CCAUUACAC AAAAGCUUCC UCAGGCGGUA AUGGGAGCCUU CCCAGUACUCC 6510 5620 6630 6640 6550 6550 6560 EUCCCAAC GGGUGGAGUA UCUCUUGAAA GCAUGGGCGG AAAAGAAGGA CCCCAUGGGU EUCCGUAUG AUACCCGAUG CUUCGACUCA ACCGUCACUG AGAGAGACAU CAGGACCGAG ACCCCAUAU ACCAGGCCUG CUCCCUGCCC GAGGAGGCCC GCACUGCCAU ACACUCGCUG 6730 6740 6750 6750 6760 6770 6780 ACCCCAUAU ACCAGGCCUG CUCCCUGCCC GAGGAGGCCC GCACUGCCAU ACACUCGCUG 6790 6800 6810 6820 6830 6840 EUGAGAGAC UUUACGUAGG AGGGCCCAUG UUCAACAGCA AGGGUCAAAC CUGCCGUUAC						
6430 6440 6450 6450 6460 6470 6480 CAMAGAMACCA 6490 6590 6510 6520 6530 6530 6540 EUGGCOUCA UCEUUUACCC UHACCUCGCC GUCGGGUCU GCGAGAAAAU GGCCCCUCUAU 6550 6560 6570 6580 6590 6690 6600 CAMUUACAC AAAAGCUUCC UCAGGCGGUA ADGGGAGCUD CCUAUGGCUU CCAGUACUCC 6510 6620 6630 6640 6650 6650 6660 EUGCCCAAC GGGUGGAGUA UCUCUUGAAA GCAUGGGCGG AAAAGAAGGA CCCCAUGGGU 6670 6680 6690 6700 6710 6720 EUUCGUAUG AUACCCGAUG CUUCGACUCA ACCGUCACUG AGAAAAACACAU CAGGACCGAG 6730 6740 6750 6760 6770 6780 EUCCGUAUAU ACCAGGCCUG CUUCGACUCA ACCGUCACUG GCACUGCCAU 6690 6600 6600 6600 6600 6600 6600 660						
### CANCAUGG CCAAAAAUGA GEUGUUCUGC GUGGACCCCG CCAAGGGGGG UAAGAAACCA ##############################	•			•		
######################################	UAAGAAACCA	CCAAGGGGGG	GUĞGACCCCG	GGUGUUCUGC	CCAAAAAUGA	ACCAUCAUGG
### ACACUCAUA AAAAGCUUCC UCAGGCGGUA ADGGRAGCUU CCUAUGGGUU CCAGUACUCC #### 6510	GGCCCUCUAU	GCGAGAAAAU	GRECEGGREA	UGACCUCGGC	UCEUUUACCC	GCUCGCCCCA
######################################	CCAGUACUCC	CCUAUGGCUU	ADGGGAGCUU	UCAGGCGGUA	AAAAGCUUCC	GACAUUACAC
TUDOSUAUS AUACCOGAUS CULTOSACUCA ACOSUCACES AGAZAGACAU CAGGACOGAS 6730 6740 6750 6760 6770 6780 ACUCCAUAU ACCAGGCCUG CUCCOUGCOC GAGGAGGCCC GCACUGCCAU ACACUCGCUG 6790 6800 6810 6820 6830 6840 TEGAGAGAC UUUACGUAGG AGGGCCCAUG UUCAACAGCA AGGGUCAAAC CUGCGGUUAC 6850 6850 6870 6880 6890 6900	CCCCAUGGGU	6553 ADDRADARA	ecadeecee	6630 UCÚCUUGAAA	6620 GGGUGGAGUA	6610 COUGCOCAAC
ACCAGECCUG CUCCCUGCCC GAGGAGGCCC GCACUGCCAU ACACUCGCUG 6790 6800 6810 6820 6830 6840 CICAGAGAGAC DUDACEUAGG AGGGCCCAUG DUCAACAGCA AGGGUCAAAC CUGCGGUUAC 6850 6850 6870 6890 6890 6900	6720 CAGGACCGAG	6710 AGAGAGACAU	6700 ACCGUCACUG	6690 CUUCGACUĆA	6680 AUACCCGAUG	6670 UUUUCGUAUG
EGAGAGAC DUDACEUAGG AGGGCCCAUG DUCAACAGCA AGGGUCAAAC CUGCGGUUAC 6850 6850 6870 6890 6890 6900	6780 ACACUCGCUG	&770 GCACUGOCAU	evecvéecce eléq	6750 CUCCCUGCCC	6740 ACCAGGCCUG	GETA GACUCCAUAU
6850 6850 6870 6890 6890 6900	6840 CUGCGGUUAC	AGGGUCAAAC	6820 UUCAACAGCA	6810 AGGCCCAUG	6800 TUUACEUAGG	6790 ACDGAGAGAC
	6900	6890	6830	6670	6850	6850

図2F

6910 GUGARAGCOC	02GCCDG	6930 CAAGGCUGCG	6940 GGGAUAGUUG	05950 06CCCACAAU	
6970 GGCGAUGACC	0960 Uaguagucau	6990 CUCAGAAAGC	7000 Cagggacug	7010 AGGAGGACGA	7020 GCGGAACCUG
7030 AGAGCCUUCA	7040 CGGAGGCCAU	. 7050 GACCAGGUAC			7080 CCCCAGACCG
7090 GAAUAUGACC	7100 UGGAGCUAAU	7 <u>11</u> 0 AAĆÁUCCIGU	7120 UCCUCAAAUG	7130 UGUCUGUGGC	7140 GUUGGGCCOG
7150	7160		7180	7190	7200
7210 Deggaaacag	7220 UUAGACACUC	7230 CCCUAUCAAU			7260 CCAGUAUGCU
7270 CCAACCAUAU	7280 GGGUUCGCAU	7290 GGUCCUAAUG	7300 ACACACUUCU	7310 UCDCCAUUCU	7320 CAUGGUCCAA
7330 GACACCCGG	7340 ACCAGAACCU	7350 CAACUUUGAG	7360 AUGUAUGGAU		7380 OGUGAAUOCU
7390 UNGGACCUUC	7400 CAGOCAUAAU	7410 UGAGAGGUUA	7420 CACGGGCUUG	7430 AOGCCHUUUC	7440 UADGCACACA
7450 UACUCUCACO	7460 ACGAACUGAC	7470 GCGGGGGCU		7490 GAAAACUDGG	7500 GGCGCCACCC
7510 CUCAGGGUGU	7520 GGAAGAGUCG	7530 GGCUCGCGCA	7540 GUCAGGGGGU	7550 CÓCUCAUCUC	7560 COGUEGAGGG
7570 AAAGOGGOOG	7580 000GGGGCCG	MADEUCUUC		UGAAGACCAA	7620 GCUCAAACUC
ACUCCAUUGC		CCUACUGGAC		eguucacogu	CEGCECCEGC
	~	7710 OGUGUOSOGC			
7750 CUCCUACUUU	7760 UCGUAGGGGU	7770 AGGCCCCUUC			
		CUCUUCCUUU	กกกดกกดกกดก		บบบบบบบบบบบบ
	COUUDUUUU	UUUUUCCCUC	nnicenecen		UUCUACUUUC
UUUCUUGGUG	GCCCCAUCUU	-7950 AGCCCUAGUC	ACGGCTUAGCT	GUGAAAGGUC	CGUGAGCCGC
0990 Acceptage	QQQ8 UDDDDUDADA	AACUGGUCUC 8010	0203 COGCAGAUC	8030 Augu	8040

図3A

			40		60
ACCCGCCCCU	AAUAGGGGCG	ACACUCCECC	AUGAAUCACU	CCCCUGUGAG	GAACUACUGU
70	AAAGCGUCUA	90	100	GUOGUAÇÃO	L20
CUUCACGCAG		GCCAUGGCGU	UAGUAUGAGU	GUOGUAÇÃO	CUCCAGGCCC
	GGAGAGCCAU	AGUGGUĆUGĆ	GGANCCGGUG		AAUUGCCGGG
AAGACUGGGU	500	AUBRACCCAC	UCUAUGOCOG	GCCAUUUGGG	cenececce
250	948	270	280	290	300
Caagacugcu	DAUDADODDA	CGUUGGGUUG	CGAAAGGCCIJ	UGUGGUACUG	CCUGAUAGGG
310	320	330	340	350	360
UGCUUGOGAG	UGCCCCGGGA	GGUCUCGUAG	ACCGUGCACC	ÁUGAGCACAA	AUCCCAAACC
370	380	390	400	AUUGAACAAG	420
UCAAAGAAAA	ACCAAAAGAA	ACACUAACOG	UCGCCCARUG		AUGGAUUGCA
430	440	450	460	470	CACARCAGAC
CGCAGGUUCU	COGGCCGCUU	GGGUGGAGAG	GCUAUUCGGC	UAUGACUGGG	
AAUCGGCUGC	ACABWARCOG 200	CCGUGUUÇCG	Ģ ÇŪĞUÇAĞÇĞ	CAGGGGGGCCC	CCCUUCUUUU
550	560	S70	580	590	CGCGGCUAUC
TGUCAAGACC	GACCUGUCOG	GUGCCCUGAA	UGAACUGCAG	GACGAGGCAG	
GDGGCDGGCC	620 ACGACGGGOG	630 UBCCUDGCGC	AGCUGUGCUC	650 GACGUUGUCA	. 660 CUĞAAGCGGG
670	680	690	700	710	720
AAGGGACUGG	CUGCUAUUGG	GCGAAGUGCC	GGGGCAGGAU	CUCCUGUCAU	CUÇACCUUĞC
730	740	750	760	770	780
UCCUGCCGAG	ADOUAUDAAA	UCAUGGCUGA	UGĆAAUGOSG	CGGCUGCAUA	CGCUUGAUCC
790	600	810	820	630	GUACUCGGAU
GGCUACCUGC	CCAUUCGACC	ACCAAGOGAA	ACAUGGCAUC	GAGCGAGCAC	
REG	·	870	. 880	. 890	900
910	920	930	940	950	OCCUGACCCA
OGAACUGUUC	CCAGGCCA	AGGOGOGOAU	GCCCGAOGGC	GAGGAUCUCG	
970	960	990	1000	1010	GAUUCAUCGA
DGGGGAUGCC	UGCUUGCGA	AUAUCAUGGU	GGAAAAUGGC	CGCUUUUCUG	
1030	1040	1050	1060	1070	1080
	CUGGGUGUGG	CGGACCGCUA	UCAGGACAUA	GOGUUGGOUA	CCOGUGADAU
1090	1100	1110	1120	1130	1140
UGCUGAAGAG		AAUGGGCDGA	CCCCUUCCUC	GUGCUUUACG	GUAUCGCCSC
1750	•	1170	1180	1190	1200
1210	•	1230	1240	1250	1260
1270	1230	1290	1300	1310	1320
	UAUGUUAUGU				
AACCUGGCCC	1340 UGUCUUCUUG	1350 DUVADBABDA ACGAGOA	1360 CUAGGGGUCU		L380 GCCAAAGGAA

€.

図3B

1440 UGAAGACAAR	1430 GGAAGCUUCU	1420 CAGUUCCUCU	1410 Gugaaggaag	1400 Guugaauguc	1390 Decaregucu
1500 AGGUGCCUCU	1490 ACCUGGCGAC	1480 GGAACCCCCC	1470 UGCAGGCAGC	. 1460 AGCGACCCUU	1450 CAACGUCUGU
1560 CAGUGCCACG	GCCACAACCC	CUGCAAAGGC	UAAGAUACAC	GCCACGUGUA	GOGGOCANAN
1620 UUCAACAAGG	1610 CUCAAGOGUA	1600 AAUGGCUCUC	1590 Gaaagaguca	0.82 <i>L</i>) GNOORUAD	L570 UUGUGAGUUG
1680 CCUCGGUGCA	1670 UGAUCUGGGG	1650 GUAUGGGAUC	1650 GUAÇÇÇAUU	1540 UGCCCAGAAG	1530 GGCUGAAGGA
1740 AACCACGGGG	AGGCCCCCCG	AAAAACGUCU	UCGAGGUUAA		CAUGCUUUAC
1800 UUACGCCCAG	1790 CCAUCACCCC	1780 ACCAUGGCCC	. 1770 ACACGAVAAU	1760 CCUUUGAAAA	1750 ACEUGGUUGU
1850 Caagacagaa	CGGGGCGUGA	GUGAGCAUGA	CUCUAUAGUG	1820 GUCUCUUGGG	1810 CAGACACGAG
1920 AACAUCCAUU	CCUUCCUCGG	GUCACUCAGU	CCUGUCCACA	AGGUCCAAGU	
1980 CGGCUCGCGG	AGACACUAGC	GCUGGCAACA	UUACCACGGA	UAUGGACUGU	UCGEGGGUCU
GCCCAGCCCU	UGGUCGEGUG	GAGGGGGACU		CGCAGADGUA	GECCCGGGCA
UUUGGUCACO	DOGRCCUGUA	UGDGGAGCGG		ADDUUGGA	CONGGGGGCCY
GCUGCUCUCC	ACCCCGGACC	CGCGGGGACA		AUGUCAUCOC	CCGAACCCUG
CCCUAGGGGC	CUGUGCUUUG	,UCGGGGGGGC		UUUCGACCUU	CCGAGACCCC
2280 GUCCADAGAU	GUGUGGCUAA	nacnanage		GAAUCUUCCG	CACGCUGUCG
2340 DGACÁACAGO	CCACLUUUAG	ACCCGGUCUC	CGACAUCGUC	UUGAGACGCU	DUCAUCCOCG
2400 CACUGGCAGU	DECACECCCC	GUGGGGUACU	GACCUAUCAG	CUGUGCCCCA	ACACCACCAG
·	GCUAUAAAGU	ecceccaee	CGDCGCGUAC	CCAYCCACC	2410 GGAAAAAGCA
2530 ACAUGGCAUC	2510 UGUCCÁAGGC	GGGGGGUACU	CCUGGGAUUU	2480 UGGCUGCICAC	2470 ANUCCCUCSG
DACADACUCC	2570 GEGAGCCCAU	2560 GUGACGACCG	2550 AGUCAGAACU	2540 UUAGGACUEG	2530 AACCCCAACA
2640 CAUCAUCAUA	Sesson Se	2620 UGCGCAGGCG	2610 CGAUGGGGGC	2500 AAUUCCUCGC	0658 Acquauggua
2700 AGUCCUTGAC	ecynceeeyc 5480	2680 ACUAUUCUOG	2670 GCAUGCUACC	2660 GCCACUCUGU	DECEAUGAAU
2760 CCCCGGGUCG	2750 CCCCACCCC	2740 GDACUGGCCA	2730 CAGGCUAACU	2720 CAGCCGGGGU	CAAGCAGAGA

図3C

2020 UGAGAUCCCC	2810 GACAGGAGGG	GUAGCCCUCG	2790 UAUAGAGGAG	2789 CCCAUGCCAA	2770 GUGACAACOC
2880 GAUUUUÜÜĞĞ	2870 GGAGGCACUU	2860 AUCAAGGGAG	2850 CCUGUCUUAC	2840 GGGGGUUVČC	· 2030
0408 CUCGAACGCUC	S930 GGGGCAUGGG	ACGCCCUUC 2920	2910 CGAGCIICECA	2900 Ahaaguguga	2690 CACUCAAGA
•	2990	. 2980	2970	2960	2950
3060	O ROSO	3040	2030	1020	2070
GAUCGACUGC 3120 UAUAACCACA	3110 CCACCITICAC	3100	3090	3080	3070
3180 GGGUAGAGGA	3170	. 3160	3150	31 <i>0</i> 0	3130
3240	3230	3220	3210	3200	APTF
GUUUGACAGU 3300	3290	3280	3270	3260	3250
DAEE	AUGAGCUCUC	GCUGCUUGGU 33A0	CGACGCAGGA 3330	GUGAGUGCUA.	GUAGUACUCU
CCAGGACCAC	USCCUGUGUG 3410	ACGCCUGGCU	GUAUUUCAAC 3398	GGCUCAGGGC	AOGACCGUCA
3480	UAGACGCUCA	CUCACACACA	UUUCACOGGC	GGGAGGCAGU	CUUGAGUUUU 3430
CACAGUGUGC	CCUAUCAGGC	UACUUAGUAG	AAAUUUOGCA	AGUCGGGGA	CAGACAAAGC
GACUCGACUC	GGAAGDGCUU	GACCELCAUGU	•	AAGCCCCCCC	3490 GCCAGGGCCA
DACCAACGAG	DEGECUCUGU	CUGUACOGUU	UNCACCUICUE	3560 DUGUGGGCCC	3550 AAGCCCACGC
3660 UGACCUCGAG	3650 GCAUGCAAGC	3640 AUCGCCACAU	3630 GACAAAAUAC	3620 CACACCCCGD	3619 AUUUUUA
3720 CGCGUAUUGC	3710 CAGCCGUOGC	3700 GGAGUCUVAG	3690 CCUGGCUGGG	3680 GCACGUGGGU	3670 GUÇAUĞACCA
3780 AGCUGUGGUG	3770 UCAACCAGCG	3760 CGUUUACACA	3750 CAUCAUUGGC	3740 GGUGUGUUUC	3730 UUAGCGACCG
3840 UGCCUCCAGA	OESE Sukassassu	3820 Duugaugaga	3810 CUAUGAGGCU	3800 AGGAGGUCCU	3790 GCUCCGGACA
3900 GAUCCAAGGC	3890 UGAAGUCAA	3880 GCCEAGAUGC	3870 GCAGCGGAVA	0880 Degrapatu	3650 GOGGCUCUCO
3960 AGCUUCGUGG	3950	3940	3930	3920	3910
4020	4010	4000	3990	3980	ትዕ ፖስ
CAUUCAGUAC 4080		•			
GGCAUUCAGC	CUUCCADGAU	CCUCCACAG	GCCAGGGAAC	DEACYVCVCA	CUCGCAGGAC
4140 UCUGGGGGC	4130 UUCUUAACAU	ACCACCAUCC	4110 GUCAACUAGO	d100 CCAGUCCGOU	4090 GCCECCCUCA

図3D

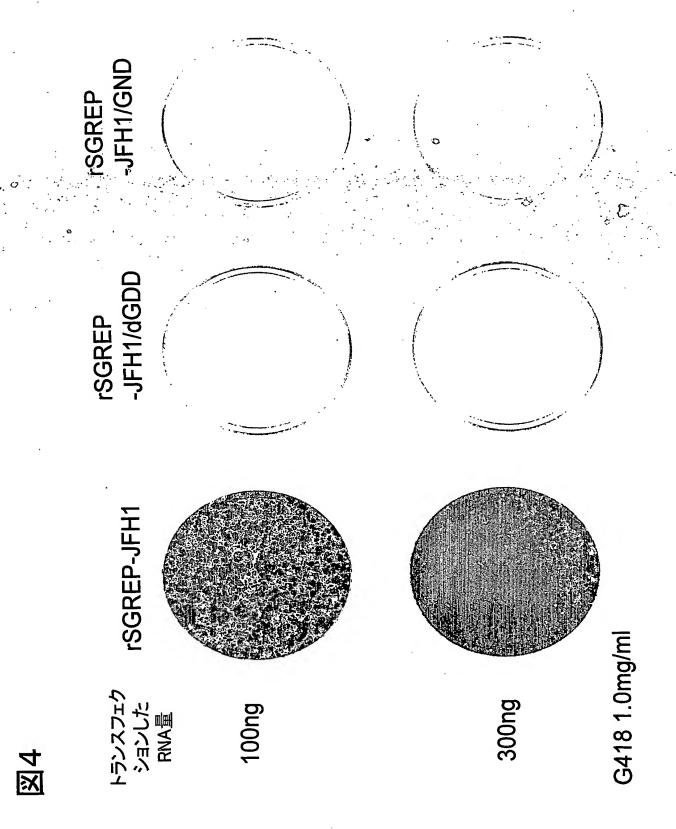
4200 CAGUGGCCUG	GCOUUGUGU	4180 GGGCCACUG	6170 6CCACCOGCG	0180 CCCAAAUUGC	4150 UGGÇUGGCGÜ
4260 CCUGGCAGGC	4250 UGGUGGACAU	4240 GGUAAAGUGC	4230 CAUAGGCIUG	4220 CUGUUGGCAG	4210 GUGGGAGCUG
4320	4310 UCAUGUCUGG	4300	4290	4260	4270
4380	4370 CUCCAGGUGC	4360	4350	4340	4330
				•	4390
4500	4490 A03U0GCCCC	4480	4470	4450	4450
dACACACOAC	4950	AGAGGAAAOC	AFRA	DECULATIONS	UGGAUGAACA
CACUAUAACU	nneachana	ACCCARCUSC	GCAGCGUGUE	CEGAUGCEUC	GUGACGGAGU
VACCECCEC	4610 GCCCCAUCOC	ACDGAGGAUU	CAACUGGAUC	GGAGACUUCA	AGUCUACUCA
4580 UAAGAACUGG	4670 UAAÇAÇACUU	4660 UGUACCAUCC	.4650 GGACUEGGUC	4640 GOGAUGUGUG	eccuescuce
4740 CCAAAAGGGG	4730 UUAUCUCUUG	4720 GGCCUCCCCU	4710 AAAGAUGCCU	4700 AGCUGUUCCC	4690 CUGACCUCCA
008 <i>&</i> 2442292999	4790 GAUGCCCCGG	4780 AUGACCACAC	4770 CACUGGCAUC	#760 VGUEEGOOGG	4750 UACAAGGGOS
4860	4850 GACCCAAAAC	4840	4630	4820	agio
4920	4910 GCCAGUGCUU	4900	4890	4880	4870
4980		4960	4950	4940	4930
5040	. 5030	5020	5010	5000	4990
	ACAACUUAAA				
UNGGUCCGCC	5090 UACAAAUCCA	GEGGACGGAG	OUUCUCUUSG	CUCCAGAGUU	CAACUCCCCU
5160 CAAUUCAUUU	5150 GCGUUGGGCU	. 5140 GUCUCGUUCA	5130 COGGGAUGAG	5120 ACCCGUUUUU	5110 CCCACACCAA
5220 GAUGUCCAUG	5210 CUGAGGUAGU	5200 GAGCCCGACA	5190 CUGUGACCCU	5180 CECAGCULICO	5170 GUCGUCGGGU
5280 GCGGGGGUCA	S270 GGCGUUVAGC	5250 GCDGCAGCGC	5250 CACGGCGGAG	5240 CAUCCCAUAU	5230 CUAACAGACC
5340	5330 CGCCAUCGCU	5320	5310	5300	0es2
5400	5390	5380	5370	5360	5350
	AUGOCAACCU				•
CCUCGACUCA	5450 UUCUGGACUC	AAAGUGGUCG	5430 GUCUGAGUCC	5420 Tuccigauaga	9410 ABUBDBBDBDBDBB
5520 GCI/CCCCAGG	5510 CGGAGUAUAU	UCAGUACCAU	3490 CCUUGAGCCU	548û Abgagagaa	9470 Augacogágo

図3E

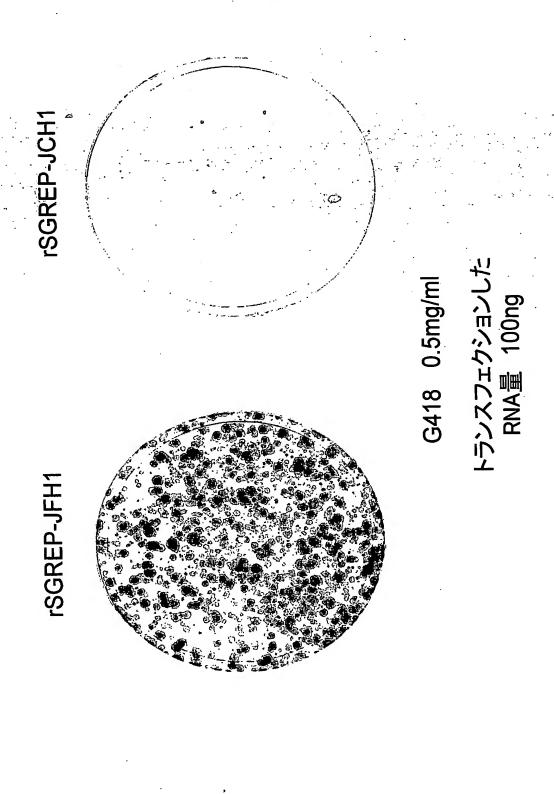
5530 AAGAGGTEYCC	5540 C200GGCCUU	5550 ACCGGCUUGG	\$560 GCGCGCCDG	5570 AUUACAACOC	SSSO ACCICCUSOSA
5590	5600	5610	5620	5630	5640
		OUACCAACCA 5670			
CCCAAAAAGA	CCCCGACGCC	UCCUCCAAGG	AGACGCCGGA	CAGUGGGUCU	GAGCGAGAGC
ACCADAGGAG	AUGCCCUCCA	5730 ACAGCUGGCC 5790	5740 AUCAAGUCCU	5750 GUGGCCAGCC	CCCCCCAAGC
5770 VEGCGAUUCAG	5780 GCCUUUCCAC	9790 GEGGGGGAC 5850	eccecceiven asúo	COGGCGAUGG	GACACCCCU
GACGAGUUGG	CUCUUUCGGA	GACAGGUUCU	ACCUCCUCCA	DECECCOCC	CEAGGGGGAG
APRA T	5900	5910 GCCUGAGCAG	5920	5930	5940
5950 GAGGCAGCUC	5960 COGGCUCGGA	5970 CUCGGGGGCC	DSGUÇÜACIDI VGGUÇÜACIDI	5990 GCUCCGAGGA	6000 GGAUGACUCC
5010 GUCGUGUGCU	6020 GCUCCAUGUC	6030 MUAUUCCUGG	6040 ACCEGGGCUC	6050 UAAUAACUCC	6060 UUGUAGCCCC
6070 GAAGAGGAAA	6080 AGUUGCCAAU	6090 UAACUCCUUG	6100 AGCAACUOGC	6110 UGUUGCGAUA	6120 CCAUAACAAG
6130 GUAUACUGUA	6140 CUACAUCAAA	6150 GAGUGCCUCA	6160 CUAAGGGCUA	6170 Aaaagguaac	6180 DOUUGAUAGG
6190 AUGCAAGUGC	6200 UCGACGCCUA	5210 UUAUGAUUCA	6220 GUCUUAAAGG	6230 ACAUCAAGCD	. 6240 AGOGGOCUCE
6250 AAGGUCAGCG	6260 CAAGGCUCCU	6270 CACCUUAGAG	6280 GAGGCGUGCC	6290 AAUUGACOOÇ	ACCCCACUCU
GCAAGAUCCA	AGUAUGGGUU	6330 DAAUCGGCUAAG	GAGGUCCGCA	GCUUGUCOGG	GAGGGCCGUC
6370 AACCACAUCA	0859 AGUCCCGUGA AGUCCCGUGA	6390 GAAGGACCUC	0400 UJGGAAGAÇUU	6410 CACAAACAOC	6420 AAUUCCUACA
6430 ACCAUCAUGG	6440 CCAAAAAUGA	6450 GGUGUUCUGC	6460 606680000	6470 CCAAGGGGGG	0848 ACCAGGGGGGU
6490 GCUCGCCUUA	6500 UCGUUUACCC	6510 DGACCUCGGC	GUCAGGGUCU 6520	OCGAGAGAD	6540 GGCCCUUUAU
6950 GAUGUCACAC	. 6560 AAAAGCUUCC	6570 UCAGGCGGUG	6580 AUGGGGGGUU	6590 CUUAUGGCUU	6600 CCAGUACUCC
6610 CCCGCUCAGC	6620 GGGUGGAGUU	0630 UCUCUUGANG	6640 GCACGGGGG	0630 ADKDRDAKKA	6660 CCCUAUGGGU
6670 SUAUQQUUUU	6680 AUACCCGÁUG	6690 CUUUGACUCA	6700 ACCGUCACUG	6710 AGAGAGACAU	6720 CAGGACUGAG
6730	6740	6750 COCCOUACOC	6760	6770	6780
5790	6800	6810	- 6820	6830	6840
acugagagag	UCDAUGUGGG	AGGGCCCAUG	OUCAACAGCA	AGGCCAGUC	CUCCECCUAC
6850 AGGCGUUGCC	6860 600022000	6870 GGUGCUUACC	0883 Devaudauca Devaudauca	6990 GGAACACCAU	690D CACAUGCUAU

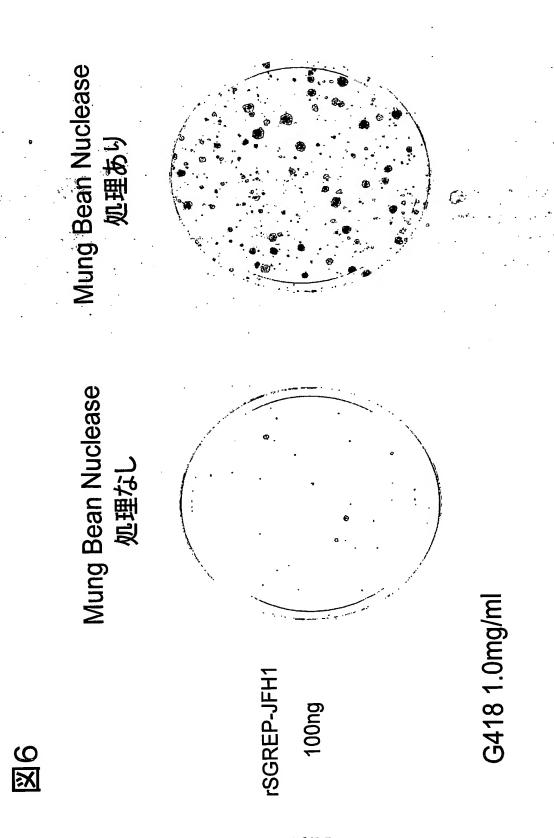
図3F

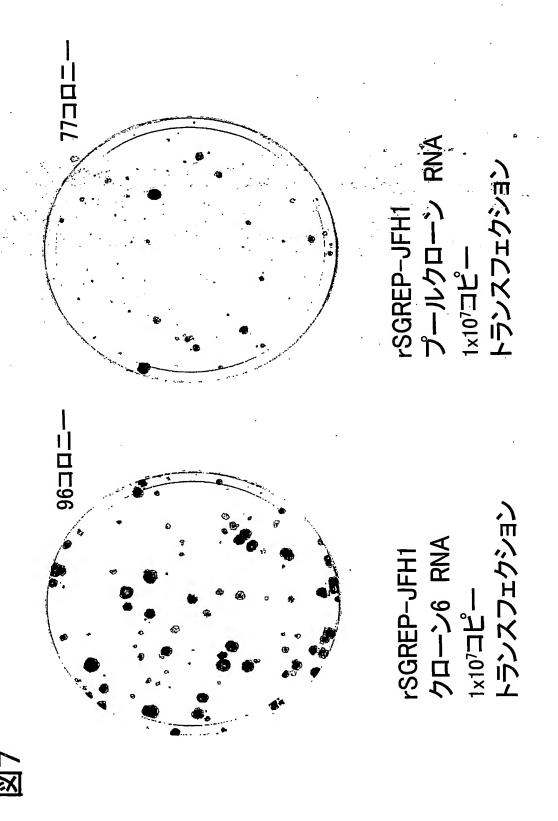
6910	6920	6930	6940	02000000000000000000000000000000000000	6960
GUAAAAGCCC	UAGCGGCUUG	CAAGGCUGCG	GGGAUAAUUG		GCUGGUAUGC
6970	6980	6990	7000	7010	
Geogrogacu	UGGUCGUCAU	CUCAGAAAGC	CAGGGGAÇUG	AGGAGGACGA	
7030	7040	7050	7060	7070	7080
agagccouca	CCCACCUAD	GAOCAGGUAU	UCUGCCCCUC	CUGGUGACCC	CCCCAGACOG
7090	7100	7110		7130	7140
GAAUAUGACC	UGGAGCUAAU	AACAUCUUGU		UGUCUGUGGC	ACUUGGCOCA
CAGEGCCCC	7166 GCAGAUACUA 7220 UUAGACACUC 7260 GCGUUCGCAU	7170 CCUGACCAGA	7180 GACCCCACCA	CIUCANUISC	ccccccccc , 1500
7310 ,UGGGAAACAG	UUAGACACUC	7230 CCCOGUCAAU	7240 UCAUGGOVEG	7250 GAAACAUCAU	CCAGUACGCU
7270 CCAACCAUAU	7280 GGZUUCGCAU	7290 GGUCCUGAUG	7300 ACACACUICU	UCUCCAUDCU	CAUGGCCCAG
7330	7340	7350	7360	7370	7360
GAÇACOCUAG	ACCAGAACCU	DARCUUUSAA	UAGGACIAUDUA	CGGUGUACUC	CGUGAGUCCU
	7400 CAGCCAUAAU		CACCGCCUTG	ACCCUUCUC	
7450	7460	7470	7460	7490	7500
UACACUCCCC	ACGAACUGAC	GCGGGGGGG	TCAGCCCUCA	GAAAACUUGG	GGOGCCACOC
	7520 GGAAGAGUOG				CCGGGGGGGG
7570	7580	7590	7600	7610	7620
AGGGCGGCCG	DUUGOGGUOG	GUACCUCUUC	AACUGGGCGG	UGAAGACCAA	GCUCARACUC
7630	7640	7650	7660	7670	7680
ACCCCCCCCCCCCCC	CGGAGGCACG	CCCCCCGGAU	UDGUCCAGUU	GGUUUACOSO	CGGCGCCGGC
GGGGGGGACA	7700 UUUAUCAÇAG	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	GCCCGACCCC	GCCUATUACU	CCUUAGCCUA
7750	7760	7770	7780	7790	7800
CUCCUACUUU	CUGUAGGGGU	AGGOCUCUUC	CUACUCCOCG	CUCGAUAGAG	CGGCACACAD
7810	7820	7830	7840	7850	7860
UAGCUAÇAÇU	CCAUAGCUAA	CUGUDCCUUD	000000000	0000000000	CUUUUUUUUUUUUUUUU
7870 000000000	7880 CUUUUUUUU	7890 0000000000	7900 UUUCOUCCCU	GCCACCOUA	•
7930	7940	7950	7960	7970	7980
UUUCUUGGUG	GCUCCAUCUU	AGCCCUAGUC	ACCCUNCCU	GUGAAAGGUC	CGUGAGCCGC
7990	0008	801D	8020	AUGU 8030	8040
AUGACUGCAG	VEAGUGCCEU	AACUGGUCUC	UCUGCAGAUC		-

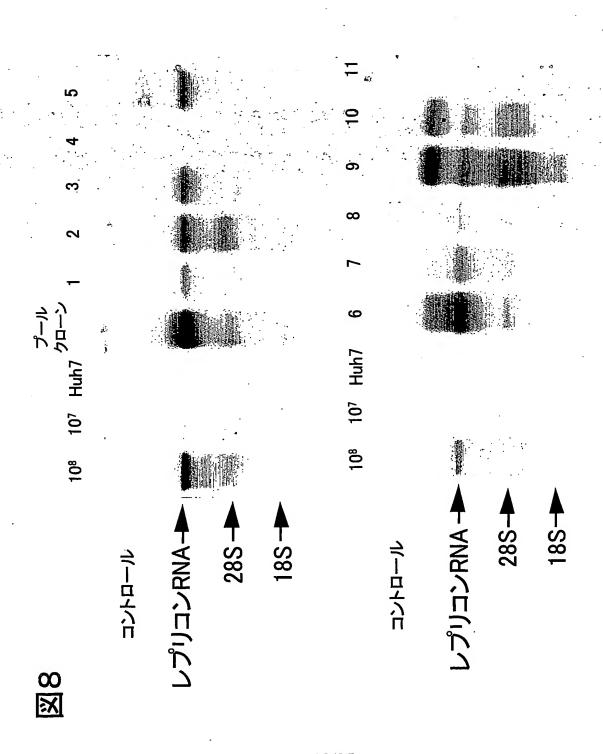


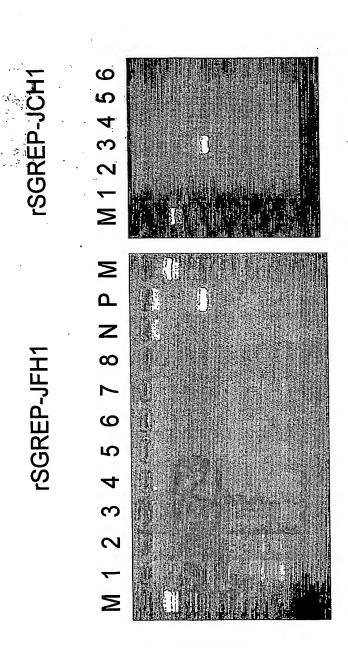
14/25



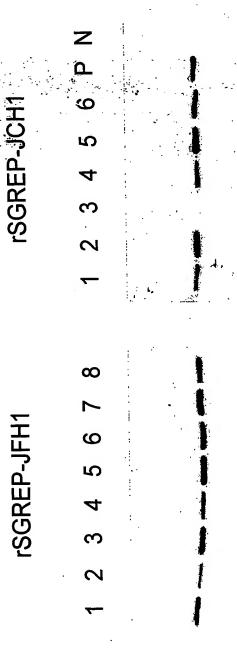


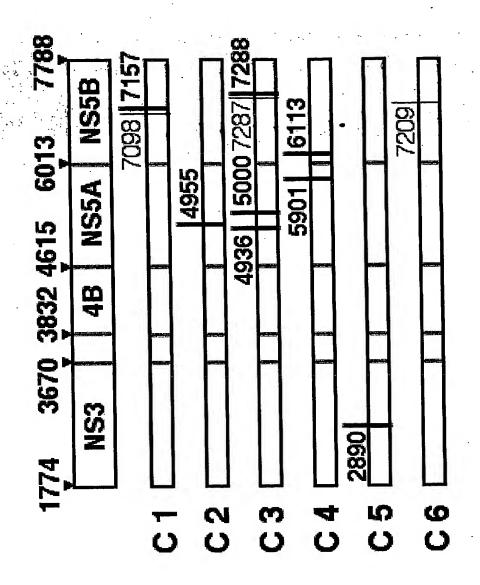


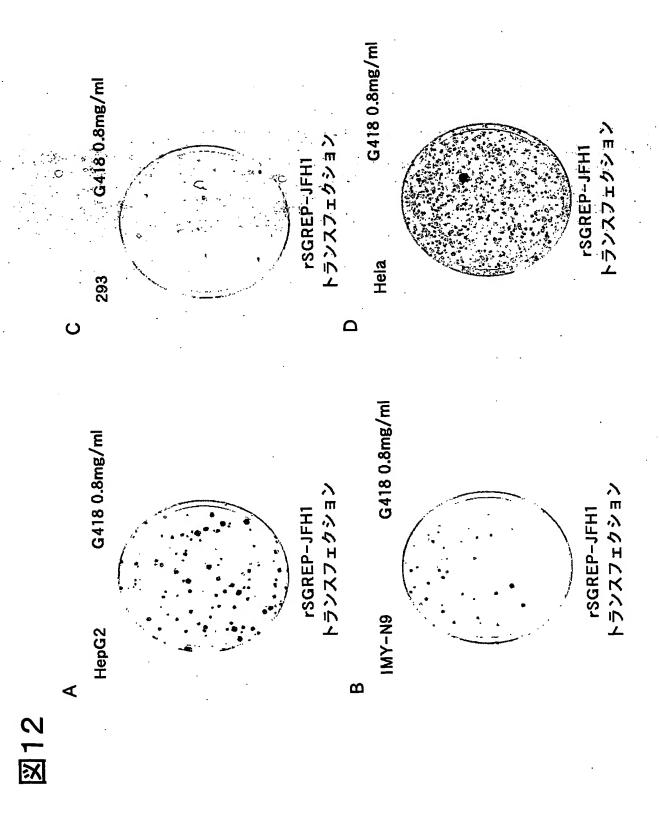




<u>図</u>



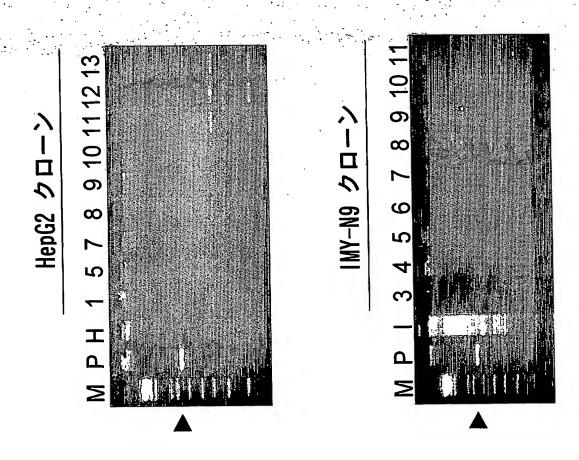


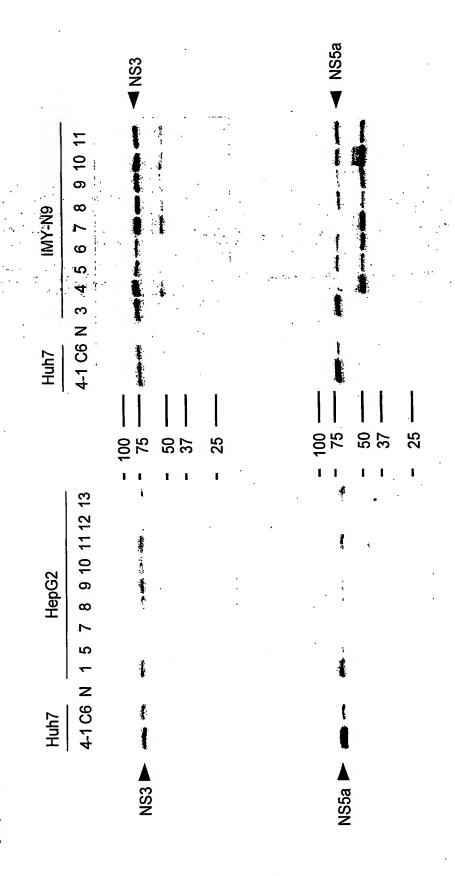


PCT/JP2003/015038.

図 1 3







SEQUENCE LISTING

<110> Toray Industries Inc.

Tokyo Metropolitan Organization for Medical Research
Johannes Gutenberg-Universitaet Mainz

<120> A nucleic acid construct comprising genomic nucleic acid of genotype 2a Hepatitis C virus, and the nucleic acid construct-transferred cells

<130> PH-1953-PCT

<140>

<141>

<150> JP 2003-148242

<151> 2003-05-26

<150> JP 2003-329115

<151> 2003-09-19

<160> 41

<170> PatentIn Ver. 2.1

<210> 1

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<220>

<223> Inventor: Wakita, Takaji

Inventor: Kato, Takanobu

Inventor: Date, Tomoko

<400> 1

accugecceu aauaggggeg acaeueegee augaaueaeu eeceugugag gaacuaeugu 60 cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360 ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggggcgcc cgguucuuuu 540 ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugeegag aaaguaueea ucauggeuga ugeaaugegg eggeugeaua egeuugauee 780 ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 uccegauucg cagegeaucg ceuucuaucg ceuucuugae gaguucuucu gaguuuaaae 1200

ccucucccuc	ccccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
cguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
caacgucugu	agcgacccuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
gcggccaaaa	egccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugccacg	1560
uugugaguug	gauaguugug'	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggcccccg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgaugau	accauggcuc	ccaucacugc	uuaugcccag	1800
caaacacgag	gccuccuggg	cgccauagug	gugaguauga	cggggcguga	caggacagaa	1860
caggccgggg	aaguccaaau	ccuguccaca	gucucucagu	ccuuccucgg	aacaaccauc	1920
ucggggguuu	uguggacugu	uuaccacgga	gcuggcaaca	agacucuagc	cggcuuacgg	1980
gguccgguca	cgcagaugua	cucgagugcu	gagggggacu	ugguaggcug	gcccagcccc	2040
ccugggacca	agucuuugga	gccgugcaag	uguggagccg	ucgaccuaua	ucuggucacg	2100
cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
ccgagaccca	uuucgaccuu	gaaggggucc	ucggggggc	cggugcucug	cccuaggggc	2220
cacgucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugcccca	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggo	acauggcauc	2520
aaucccaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
acauauggca	a aauuucucgc	cgaugggggc	ugcgcuagcg	gegeeuauga	caucaucaua	2640
ugcgaugaau	ı gecaegeugu	ggaugcuacc	uccauucucg	gcaucggaac	gguccuugau	2700
caagcagaga	a cagccggggu	cagacuaacu	gugcuggcua	cggccacaco	cccggguca	2760
gugacaacco	cccaucccga	uauagaagag	guaggecueg	g ggcgggaggg	g ugagaucccc	2820
uucuauggga	a gggcgauucc	ccuauccugo	aucaagggag	g ggagacaccı	ı gauuuucugc	2880
cacucaaaga	a aaaaguguga	cgagcucgcg	geggeeeuue	ggggcauggg	g cuugaaugco	2940

guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000 gucgccaccg acgcccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060 aauguagegg ucaeccaage uguegaeuue ageeuggaee eeaeeuueae uauaaccaea 3120 cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180 agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240 guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300 accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360 cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucuc 3420 caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480 gccagagcca aggccccucc cccguccugg gacgccaugu ggaagugccu ggcccgacuc 3540 aagccuacgc uugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600 gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660 gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720 cuggegacug gaugeguuuc caucauegge egeuugeaeg ucaaceageg aguegueguu 3780 gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840 gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900 uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau caugggaggc 4140 ugguuagcgu cccagaucgc accacccgcg ggggccaccg gcuuugucgu caguggccug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg cgcgguccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620 uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680

cugaccucua	aauuguuccc	caagcugccc	ggccuccccu	ucaucucuug	ucaaaagggg	4740
uacaagggug	ugugggccgg	cacuggcauc	augaccacgc	gcugcccuug	cggcgccaac	4800
aucucuggca	auguccgccu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
				gccagugcgc		
cccacgaacu	acaagaccgc	caucuggagg	guggcggccu	cggaguacgc	ggaggugacg	4980
				acaaucugaa		
caacuaccuu	cuccagaguu	uuucuccugg	guggacggug	ugcagaucca	uagguuugca	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguucu	gcguugggcu	uaauuccuau	5160
gcugucgggu	cccagcuucc	cugugaaccu	gagcccgacg	cagacguauu	gagguccaug	5220
cuaacagauc	cgccccacau	cacggcggag	acugcggcgc	ggcgcuuggc	acggggauca	5280
ccuccaucug	aggcgagcuc	cucagugagc	cagcuaucag	caccgucgcu	gcgggccacc	5340
ugcaccacco	acagcaacac	cuaugacgug	gacauggucg	augccaaccu	gcucauggag	5400
ggcggugugg	cucagacaga	gccugagucc	agggugcccg	uucuggacuu	ucucgagcca	5460
auggccgagg	g aagagagcga	ccuugagccc	ucaauaccau	cggagugcau	gcucccagg	5520
agcggguuuc	cacgggccuu	accggcuugg	gcacggccug	acuacaaccc	gccgcucgug	5580
gaaucgugga	ggaggccaga	uuaccaaccg	cccaccguug	cugguugugc	ucucccccc	5640
cccaagaagg	g ccccgacgcc	uccccaagg	agacgccgga	cagugggucu	gagcgagagc	5700
accauaucag	g aagcccucca	gcaacuggcc	aucaagaccu	uuggccagcc	cccucgagc	5760
ggugaugcag	g gcucguccac	gggggcgggc	gccgccgaau	ccggcggucc	gacguccccu	5820
ggugagccg	g cccccucaga	gacagguucc	gccuccucua	ugccccccu	cgagggggag	5880
ccuggagau	c cggaccugga	gucugaucag	guagagcuuc	aaccuccccc	ccaggggggg	5940
gggguagcu	c ccgguucggg	cucggggucu	uggucuacuu	ı gcuccgagga	ggacgauacc	6000
accgugugc	u gcuccauguc	auacuccugg	accggggcuc	uaauaacucc	cuguagecee	6060
gaagaggaa	a aguugccaau	caacccuuug	gaguaacucgo	uguugcgaua	ccauaacaag	6120
guguacugu	a caacaucaaa	gagogocuca	a cagagggcua	a aaaagguaac	uuuugacagg	6180
acgcaagug	c ucgacgccca	uuaugacuca	a gucuuaaagg	g acaucaagcu	agcggcuucc	6240
aaggucagc	g caaggcucci	ı caccuuggag	g gaggcgugc	c aguugacucc	accccauucu	6300
gcaagaucc	a aguauggauı	ı cggggccaag	g gagguccgca	a gcuuguccgg	gagggccguu	6360
aaccacauc	a aguccgugug	g gaaggaccu	cuggaagac	c cacaaacaco	aauucccaca	6420

accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaagaaacca	6480
gcucgccuca	ucguuuaccc	ugaccucggc	guccgggucu	gcgagaaaau	ggcccucuau	6540
gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	6600
ccugcccaac	ggguggagua	ucucuugaaa	gcaugggcgg	aaaagaagga	ccccaugggu	6660
uuuucguaug	auacccgaug	cuucgacuca	accgucacug	agagagacau	caggaccgag	6720
gaguccauau	accaggccug	cucccugccc	gaggaggccc	gcacugccau	acacucgcug	6780
acugagagac	uuuacguagg	agggcccaug	uucaacagca	agggucaaac	cugcgguuac	6840
agacguugcc	gcgccagcgg	ggugcuaacc	acuagcaugg	guaacaccau	cacaugcuau	6900
gugaaagccc	uagcggccug	caaggcugcg	gggauaguug	cgcccacaau	gcugguaugc	6960
ggcgaugacc	uaguagucau	cucagaaagc	caggggacug	aggaggacga	gcggaaccug	7020
agagccuuca	cggaggccau	gaccagguac	ucugccccuc	cuggugaucc	ccccagaccg	7080
gaauaugacc	uggagcuaau	aacauccugu	uccucaaaug	ugucuguggc	guugggcccg	7140
cggggccgcc	gcagauacua	ccugaccaga	gacccaacca	cuccacucgc	ccgggcugcc	7200
ugggaaacag	uuagacacuc	cccuaucaau	ucauggcugg	gaaacaucau	ccaguaugcu	7260
ccaaccauau	ggguucgcau	gguccuaaug	acacacuucu	ucuccauucu	caugguccaa	7320
gacacccugg	accagaaccu	caacuuugag	auguauggau	caguauacuc	cgugaauccu	7380
uuggaccuuc	cagccauaau	ugagagguua	cacgggcuug	acgccuuuuc	uaugcacaca	7440
uacucucacc	acgaacugac	gcggguggcu	ucagcccuca	gaaaacuugg	ggcgccaccc	7500
cucagggugu	ggaagagucg	ggcucgcgca	gucagggcgu	cccucaucuc	ccguggaggg	7560
aaagcggccg	uuugcggccg	auaucucuuc	aauugggcgg	ugaagaccaa	gcucaaacuc	7620
acuccauugc	cggaggcgcg	ccuacuggac	uuauccaguu	gguucaccgu	cggcgccggc	7680
gggggcgaca	uuuuucacag	cgugucgcgc	gcccgacccc	gcucauuacu	cuucggccua	7740
cuccuacuuu	ucguaggggu	aggccucuuc	cuacuccccg	cucgguagag	cggcacacac	7800
uagguacacu	ccauagcuaa	cuguuccuuu	uuuuuuuuu	uuuuuuuuu	uuuuuuuuu	7860
นนนนนนนนน	cuuuuuuuu	uuuuucccuc	uuucuucccu	ucucaucuua	uucuacuuuc	7920
uuucuuggug	gcuccaucuu	agcccuaguc	acggcuagcu	gugaaaggud	cgugagccgc	7980
augacugcag	g agagugccgu	aacuggucuc	ucugcagauc	augu		8024

<210> 2

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

(223) Description of Artificial Sequence: replicon

<400> 2

accegececu aauaggggeg acacucegee augaaucacu ececugugag gaacuacugu 60 cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 cccccuccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacugggu ccuuucuugg auaaacccac ucuaugcceg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 ugcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa aucccaaacc 360 ucaaagaaaa accaaaagaa acacuaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggggcgcc cgguucuuuu 540 ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugeegag aaaguaucea ucauggeuga ugeaaugegg eggeugeaua egeuugauee 780 ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200

3

٠ (ccucucccuc	cccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
•	eguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
ä	accuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uucccucuc	gccaaaggaa	1380
1	ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
•	caacgucugu	agcgacccuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
1	gcggccaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugccacg	1560
1	uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
	ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
,	caugcuuuac	auĝuguuuag	ucgagguuaa	aaaaacgucu	aggccccccg	aaccacgggg	1740
	acgugguuuu	ccuuugaaaa	acacgauaau	accauggccc	ccaucaccgc	uuacgcccag	1800
	cagacacgag	gucucuuggg	cucuauagug	gugagcauga	cggggcguga	caagacagaa	1860
	caggccgggg	agguccaagu	ccuguccaca	gucacucagu	ccuuccucgg	aacauccauu	1920
	ucgggggucu	uauggacugu	uuaccacgga	gcuggcaaca	agacacuagc	cggcucgcgg	1980
	ggcccgguca	cgcagaugua	cucgagcgcc	gagggggacù	uggucgggug	gcccagcccu	2040
	ccugggacca	aaucuuugga	gccguguacg	uguggagcgg	ucgaccugua	uuuggucacg	2100
	cggaacgcug	augucauccc	ggcucgaaga	cgcggggaca	agcggggagc	gcugcucucc	2160
	ccgagacccc	uuucgaccuu	gaaggggucc	ucggggggac	cugugcuuug	cccuaggggc	2220
	cacgcugucg	gaaucuuccg	ggcagcugug	ugcucucggg	guguggcuaa	guccauagau	2280
	uucauccccg	uugagacgcu	cgacaucguc	acgcggucuc	ccaccuuuag	ugacaacagc	2340
	acaccaccag	cugugececa	gaccuaucag	gugggguacu	ugcacgcccc	cacuggcagu	2400
	ggaaaaagca	ccaagguccc	cgucgcguac	gccgcccagg	gguauaaagu	gcuggugcuc	2460
	aaucccucgg	uggcugccac	ccugggauuu	ggggcguacu	uguccaaggo	acauggcauc	2520
	aaccccaaca	uuaggacugg	agucagaacu	gugacgaccg	gggagcccau	uacauacucc	2580
	acguauggua	aauuccucgo	cgauggggg	ugcgcaggcg	gegeeuauga	caucaucaua	2640
	ugcgaugaau	ı gccacucugu	ggaugcuaco	acuauucucg	gcaucgggad	aguccuugac	2700
	caagcagaga	a cagccggggu	caggcuaacı	guacuggcca	cggccacgc	ccccgggucg	2760
	gugacaacco	cccaucccaa	uauagaggag	guageceuce	g gacaggaggg	g ugagaucccc	2820
•	uucuauggga	a gggcguuuco	ccugucuua	aucaagggag	g ggaggcacul	ı gauuuucugc	2880
	cacucaaaga	a aaaaguguga	cgagcucgc	a acggcccuud	ggggcaugg	g cuugaacgcu	2940

guggcauauu	acagaggguu	ggacgucucc	auaauaccaa	cucaaggaga	uguggugguc	3000
guugccaccg	acgcccucau	gacgggguau	acuggagacu	uugacuccgu	gaucgacugc	3060
aacguagcgg	ucacccaggc	cguagacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
cagacugucc	cgcaagacgc	ugucucacgu	agucagcgcc	gagggcgcac	ggguagagga	3180
agacugggca	uuuauaggua	uguuuccacu	ggugagcgag	ccucaggaau	guuugacagu	3240
guaguacucu	gugagụgcua	cgacgcagga	gcugcuuggu	augagcucuc	accaguggag	3300
acgaccguca	ggcucagggc	guauuucaac	acgccuggcu	ugccugugug	ccaggaccac	3360
cuugaguuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgcuca	uuuccuuucc	3420
cagacaaagc	agucggggga	aaauuucgca	uacuuaguag	ccuaucaggc	cacagugugc	3480
gccagggcca	aagcgccccc	cccguccugg	gacgucaugu	ggaagugcuu	gacucgacuc	3540
aagcccacgc	uugugggccc	uacaccucuc	cuguaccguu	ugggcucugu	uaccaacgag	3600
gucacccuua	cacaccccgu	gacaaaauac	aucgccacau	gcaugcaagc	úgaccucgag	3660
gucaugacca	gcacgugggu	ccuggcuggg	ggagucuuag	cagccgucgc	cgcguauugc	3720
uuagcgaccg	gguguguuuc	caucauuggc	cguuuacaca	ucaaccagcg	agcugucguc	3780
gcuccggaca	aggagguccu	cuaugaggcu	uuugaugaga	uggaggaaug	ugccuccaga	3840
gcggcucucc	uugaagaggg	gcagcggaua	gccgagaugc	ugaagucċaa	gauccaaggc	3900
uuauugcagc	aagccucuaa	acaggcccag	gacauacaac	ccgcugugca	agcuucgugg	3960
cccaagaugg	agcaauucug	ggccaaacau	auguggaacu	ucauaagcgg	cauucaguac	4020
cucgcaggac	ugucaacacu	gccagggaac	ccugcugugg	cuuccaugau	ggcauucagc	4080
gccgcccuca	ccaguccguu	gucaacuagc	accaccaucc	uucuuaacau	ucuggggggc	4140
uggcuggcgu	cccaaauugc	gccacccgcg	ggggccacug	gcuuuguugu	caguggccug	4200
gugggagcug	cuguuggcag	cauaggcuug	gguaaagugc	ugguggacau	ccuggcaggg	4260
uauggugcgg	gcauuucggg	ggcccucguc	gcguuuaaga	ucaugucugg	cgagaagccc	4320
uccauggagg	augucaucaa	cuugcugccu	gggauucugu	cuccaggugc	ucugguggug	4380
ggagucaucu	ı gcgcggccau	ucugcgccgc	caugugggac	cgggggaagg	cgcgguccaa	4440
uggaugaaca	ggcuuaucgc	cuucgcuucc	agaggaaacc	acgucgcccc	uacucacuac	4500
gugacggagu	ı cggaugcguc	gcagcguguc	acccaacugo	uuggcucucu	cacuauaacu	4560
agucuacuca	a ggagacuuca	caacuggauc	acugaggauu	gccccauccc	augcgccggc	4620
ucguggcuc	gcgaugugug	ggacuggguc	uguaccaucc	uaacagacuu	uaagaacugg	4680

•	cugaccucca	agcuguuccc	aaagaugccu	ggccuccccu	uuaucucuug	ccaaaagggg	4740
1	uacaagggcg	ugugggccgg	cacuggcauc	augaccacac	gaugeceeug	cggcgccaac	4800
	aucucuggca	acguccgcuu	gggcucuaug	agaaucacag	gacccaaaac	cugcaugaac	4860
	accuggcagg	ggaccuuucc	uaucaauugu	uauacagaag	gccagugcuu	gccgaaaccc	4920
	gcguuaaacu	ucaagaccgc	caucuggaga	guggcggccu	cagaguacgc	ggaagugacg	4980
	cagcacggau	cauaugccua	uauaacaggg	cugaccacug	acaacuuaaa	agucccuugc	5040
	caacuccccu	cuccagaguu	uuucucuugg	guggacggag	uacaaaucca	uagguccgcc.	5100
	cccacaccaa	agccguuuuu	ccgggaugag	gucucguuca	gcguugggcu	caauucauuu	5160
	gucgucgggu	cucagcuucc	cugugacccu	gagcccgaca	cugagguagu	gauguccaug	5220
	cuaacagacc	caucccauau	cacggcggag	gcugcagcgc	ggcguuuagc	gcggggguca	5280
	ccccaucug	aggcaagcuc	cucagcgagc	cagcugucgg	cgccaucgcu	gcgagccacc	5340
	ugcaccaccc	acgguaggac	cuaugaugug	gacauggugg	augccaaccu	guucaugggg	5400
	ggcggcguga	uucggauaga	gucugagucc	aaaguggucg	uucuggacuc	ccucgacuca	5460
	augaccgagg	aagagggcga	ccuugagccu	ucaguaccau	cggaguauau	gcuccccagg	5520
	aagagguucc	caccggccuu	accggcuugg	gcgcggccug	auuacaaccc	accgcuugug	5580
	gaaucgugga	agaggccaga	uuaccaacca	cccacuguug	cgggcugugc	ucucccccc	5640
	cccaaaaaga	ccccgacgcc	uccuccaagg	agacgccgga	cagugggucu	gagcgagagc	5700
	accauaggag	augeceucea	acagcuggcc	aucaaguccu	uuggccagcc	cccccaagc	5760
	ggcgauucag	gccuuuccac	gggggcggac	gccgccgacu	ccggcgaucg	gacacccccu	5820
	gacgaguugg	cucuuucgga	gacagguucu	accuccucca	ugccccccu	cgagggggag	5880
	ccuggggacc	cagaccugga	gccugagcag	guagagcuuc	aaccuccucc	ccaggggggg	5940
	gaggcagcuc	ccggcucgga	cucggggucc	uggucuacuu	gcuccgagga	ggaugacucc	6000
	gucgugugcu	ı gcuccauguc	auauuccugg	accggggcuc	uaauaacucc	uuguagcccc	6060
	gaagaggaaa	a aguugccaau	uaacuccuug	gagcaacucgc	uguugcgaua	ccauaacaag	6120
	guauacugua	a cuacaucaaa	gagugccuca	cuaagggcua	aaaagguaac	uuuugauagg	6180
	augcaagugo	c ucgacgecua	uuaugauuca	gucuuaaagg	g acaucaagcu	agcggccucc	6240
	aaggucagc	g caaggcucci	caccuuagag	g gaggcgugco	aauugaccc	accccacucu	6300
	gcaagaucca	a aguaugggui	ı uggggcuaag	g gagguccgca	a gcuuguccgg	g gagggccguc	6360
	aaccacauca	a aguccgugug	g gaaggaccu	uuggaagacı	ı cacaaacac	aauuccuaca	6420

accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaaaaaacca	6480
gcucgccuua	ucguuuaccc	ugaccucggc	gucagggucu	gcgagaagau	ggcccuuuau	6540
gaugucacac	aaaagcuucc	ucaggcggug	augggggcuu	cuuauggcuu	ccaguacucc	6600
cccgcucagc	ggguggaguu	ucucuugaag	gcaugggcgg	aaaagagaga	cccuaugggu	6660
uuuucguaug	auacccgaug	cuuugacuca	accgucacug	agagagacau	caggacugag	6720
gaguccauau	accaggccug	cuccuuaccc	gaggaggccc	gaacugccau	acacucgcug	6780
acugagagac	ucuauguggg	agggcccaug	uucaacagca	agggccaguc	cugcggguac	6840
aggcguugcc	gcgccagcgg	ggugcuuacc	acuaguaugg	ggaacaccau	cacaugcuau	6900
guaaaagccc	uagcggcuug	caaggcugcg	gggauaauug	cgcccacgau	gcugguaugc	6960
ggcgacgacu	uggucgucau	cucagaaagc	caggggacug	aggaggacga	gcggaaccug	7020
agagccuuca	cggaggcuau	gaccagguau	ucugccccuc	cuggugaccc	ccccagaccg	7080
gaauaugacc	uggagcuaau	aacaucuugu	uccucaaacg	ugucuguggc	acuuggccca	7140
cagggccgcc	gcagauacua	ccugaccaga	gaccccacca	cuucaauugc	ccgggcugcc	7200
ugggaaacag	uuagacacuc	cccugucaau	ucauggcugg	gaaacaucau	ccaguacgcu	7260
ccaaccauau	ggguucgcau	gguccugaug	acacacuucu	ucuccauucu	cauggcccag	7320
gacacccuag	accagaaccu	uaacuuugaa	auguacggau	cgguguacuc	cgugaguccu	7380
cuggaccucc	cagccauaau	ugaaagguua	cacgggcuug	acgccuucuc	ucugcacaca	7440
uacacucccc	acgaacugac	gcggguggcu	ucageceuca	gaaaacuugg	ggcgccaccc	7500
cucagagcgu	ggaagagucg	ggcgcgugca	guuagggcgu	cccucaucuc	ccgugggggg	7560
agggcggccg	uuugcggucg	guaccucuuc	aacugggcgg	ugaagaccaa	gcucaaacuc	7620
acuccuuugc	cggaggcacg	ccuccuggau	uuguccaguu	gguuuaccgu	cggcgccggc	7680
gggggcgaca	uuuaucacag	cgugucgcgu	gcccgacccc	gccuauuacu	ccuuagccua	7740
cuccuacuuu	cuguaggggu	aggccucuuc	cuacuccccg	cucgauagag	cggcacacau	7800
uagcuacacu	ı ccauagcuaa	cuguuccuuu	นนนนนนนนน	เบนนนนนนนนน	นนนนนนนนน	7860
นนนนนนนนน	cuuuuuuuuu	uuuuucccuc	uuucuucccu	ucucaucuua	uucuacuuuc	7920
uuucuuggug	g gcuccaucuu	agcccuaguc	acggcuagcu	gugaaaggud	cgugagccgc	7980
augacugcag	g agagugccgu	aacuggucuc	ucugcagauc	augu		8024

<210> 3

<211> 9678

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341).. (9442)

<400> 3

acctgccct aatagggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60
cttcacgcag aaagcgccta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120
ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180
aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgcccccg 240
caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300
cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct 355
Met Ser Thr Asn Pro
1 5

aaa cct caa aga aaa acc aaa aga aac acc aac cgt cgc cca gaa gac 403
Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Glu Asp
10 15 20

gtt aag ttc ccg ggc ggc ggc cag atc gtt ggc gga gta tac ttg ttg 451
12/117

Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val	Gly	Gly	Val	Tyr	Leu	Leu	
			25					30					35			
ccg	cgc	agg	ggc	ccc	agg	ttg	ggt	gtg	cgc	acg	aca	agg	aaa	act	tcg	499
Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	Thr	Thr	Arg	Lys	Thr	Ser	
		40					45					50				
• .				•			:		·	٠.		. :		, .		S
gag	cgg	tcc	cag	cca	cgt	ggg	aga	cgc	cag	ccc	atc	ccc	aaa	gat	cgg	547
			•• •		٠.	·		٠.		-		Pro				
	55				. : .	60			· · · · · ·		65		•			•
•						. •						,	•		P	•
cgc	tcc	act	ggc	aag	gcc	tgg	gga	aaa	cca	ggt	cgc	ccc	tgg	ccc	cta	595
												Pro				•
70				•	75	-	-	-	•	80					85	•
tat	ggg	aat	gag	gga	ctc	ggc	tgg	gca	gga	tgg	ctc	ctg	tcc	ccc	cga	643
												Leu				
				90					95					100		
										•						
ggc	tct	cgc	ccc	tcc	tgg	ggc	ccc	act	gac	ccc	cgg	cat	agg	tcg	cgc	691
Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro	Arg	His	Arg	Ser	Arg	
			105	j				110					115			
aac	gtg	ggt	aaa	gto	ato	gac	acc	cta	acg	tgt	ggc	ttt	gcc	gac	ctc	739
															Leu	
		120					125					130				
ats	ggg	, tac	ato	ccc	gto	gta	ggc	gcc	ccg	ctt	agt:	ggc	gcc	gcc	aga	78
															Arg	

PCT/JP2003/015038

135

140

145

gct	gtc	gcg	cac	ggc	gtg	aga	gtc	ctg	gag	gac	ggg	gtt	aat	tat	gca	835
Ala	Val	Ala	His	Gly	Val	Arg	Val	Leu	G1u	Asp	Gly	Val	Asn	Tyr	Ala	
150					155					160					165	
aca	ggg	aac	cta	ccc	ggt	ttc	ccc	ttt	tct	atc	ttc`	ttg	ctġ	gcc	ctg	883
				•	•		Pro			•	٠.		•		•	. <i>i</i> .
		•	. •	170	• •				175	• • • • • • • • • • • • • • • • • • • •		÷ _		180		:
<i>.</i>				÷: ,	• • •	•		•						:		
ttg	tcc	tgc	atc	acc	gtt	ccg	gtc	tct	gct	gcc	cag	gtg	aag	aat	acc	931
							Val									
		-	185			•		190	•				195			
agt	agc	agc	tac	atg	gtg	acc	aat	gac	tgc	tcc	aat	gac	agc	atc	act	979
							Asn									
		200					205					210				•
					•											
tgg	cag	ctc	gag	gct	gcg	gtt	ctc	cac	gtc	ccc	ggg	tgc	gtc	ccg	tgc	1027
															Cys	
	215					220			-		225					
gag	aga	gtg	ggg	aat	acg	; tca	cgg	tgt	tgg	gtg	cca	gto	tce	, cca	aac	1075
															Asn	
230					235					240					245	
ate	gct	gts	g cgs	g cag	, ccc	ggt	gcc	cto	ace	g cag	g ggt	cte	g cgg	g ac	g cac	1123
															r His	
				250		.			258					260		
				500												

atc	gat	atg	gtt	gtg	atg	tcc	gcc	acc	ttc	tgc	tct	gct	ctc	tac	gtg	1171
Ile	Asp	Met	Val	Val	Met	Ser	Ala	Thr	Phe	Cys	Ser	Ala	Leu	Tyr	Val	
			265					270					275			
•															٠	
ggg	gac	ctc	tgt	ggc	ggg	gtg	atg	ctc	gcg	gcc	cag	gtg	ttc	atc	gtc	1219
G1y	Asp	Leu	Cys	Gly.	Gly	Val	Met	Leu	Ala	Ala	G1n	Val	Phe	Ile	Val	
•		280			•		285				-	290			•	• •
				*	.:				×:		•					
tçg	ccg	cag	tac	cac	tgg	ttt	gtg	caa	gaa	tgc	aat	tgc	tcc	atc	tac	1267
Ser	Pŗo	Gln	Tyr	His	Trp	Phe	Val	Gln	Glu	Cys	Asn	Cys	Ser	Ile	Tyr	
	295					300					305					
cct	ggc	acc	atc	act	gga	cac	cgc	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	G1y	Thr	Ile	Thr	Gly	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310					315					320					325	
tgg	tcg	ccc	acg	gcc	acc	atg	atc	ctg	gcg	tac	gtg	atg	cgc	gtc	ссс	1363
Trp	Ser	Pro	Thr	Ala	Thr	Met	Ile	Leu	Ala	Tyr	Val	Met	Arg	Val	Pro	
				330					335					340		
gag	gtc	atc	ata	gac	atc	gtt	agc	ggg	gct	cac	tgg	ggc	gtc	atg	ttc	1411
Glu	Val	Ile	Ile	Asp	Ile	Val	Ser	Gly	Ala	His	Trp	G1y	Val	Met	Phe	
			345					350					355			
ggc	ttg	gcc	tac	ttc	tct	atg	cag	gga	gcg	tgg	gcg	aag	gtc	att	gtc	1459
						Met										•
-		360	,				365	•		•		370				

í	atc	ctt	ctg	ctg	gcc	gct	ggg	gtg	gac	gcg	ggc	acc	acc	acc	gtt	gga	1507
	[le	Leu	Leu	Leu	Ala	Ala	Gly	Val	Asp	Ala	Gly	Thr	Thr	Thr	Val	Gly	
		375					380					385					·
											•						•
	ggc	gct	gtt	gca	cgt	tcc	acc	aac	gtg	att	gcc	ggc	gtg	ttc	agc	cat	1555
	G1y	Ala	Val	Ala	Arg	Ser	Thr	Asn	Val	Ile	Ala	G1y	Val	Phe	Ser	His	
	390	•			ta vy	395	-	<i>:</i>	. محر د د د د		400	•				405	,
						. 1,9 2 . • .		, .		, = () ·		• • • •			٠	-	
	ggC	cct	cag	cag	aac	att	cag	ctc	att.	aac	acc	aac	ggc	agt	tgg	cac	1603
				-	•							•				His	
	U			•	410				•	415			•		420		
									•			•					
	atc	aac	cgt	: act	gcc	ttg	aat	tgc	aat	gac	tcc	ttg	aac	acc	ggc	ttt	1651
					_				Asn								
				425					430					435			
	cto	gcg	g gc	c tt	g tto	tac	acc	aac	cgc	ttt	aac	tcg	tca	ggg	tgt	cca	1699
								•								s Pro	
			44					445					450				
	ggg	g cg	c ct	g tc	c gco	tgo	c cgc	aac	atc	gag	get	tto	cgg	g ata	ggg	g tgg	1747
	Gly	y Ar	g Le	u Se	r Ala	a Cys	s Arg	g Asr	ı Ile	Glu	ı Ala	Phe	e Arg	; Ile	e G1	y Trp	
		45	5				460)				46	5				
														•			
	gg	c ac	c ct	a ca	g ta	c ga	g ga	t aa	t gto	aco	aat	t cc	a gag	g gat	t at	g agg	1795
						•										t Arg	
	47					47					480					485	•
	••	-															
	cc	g ta	c te	gc tg	g ca	c ta	c cc	c cc	a aag	g cc	g tg	t gg	c gt	a gt	c cc	c gcg	1843

Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys Gly Val Val Pro Ala 490 495 500 agg tot gtg tgt ggc cca gtg tac tgt ttc acc ccc agc ccg gta gta 1891 Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val 515 505 gtg ggc acg acc gac aga cgt gga gtg ccc acc tac aca tgg gga gag Val Gly Thr Thr Asp Arg Gly Val Pro Thr Tyr Thr Trp Gly Glu 530 **...520** 525 aat gag aca gat gtc ttc cta ctg aac agc acc cga ccg ccg ggc 1987 Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Gln Gly 535 540 545 tca tgg ttc ggc tgc acg tgg atg aac tcc act ggt ttc acc aag act 2035 Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr 550 555 560 565 tgt ggc gcg cca cct tgc cgc acc aga gct gac ttc aac gcc agc acg 2083 Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Ala Ser Thr 570 575 580 gac ttg ttg tgc cct acg gat tgt ttt agg aag cat cct gat gcc act 2131 Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Asp Ala Thr 585 590 595 tat att aag tgt ggt tct ggg ccc tgg ctc aca cca aag tgc ctg gtc 2179 Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val

600

605

610

				•			•									
cac	tac	cct	tac	aga	ctc	tgg	cat	tac	ccc	tgc	aca	gtc	aat	ttt	acc	2227
His	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Cys	Thr	Val	Asn	Phe	Thr	
	615					620					625					
0+0	++5.	224	ata	202	at a	tat	σta	 	σσσ	øt.t.	ฮลฮ	cac	agg	ctc	acg	2275
<i>:</i>		-								•						1
•	Phe	Lys	TIE	Arg	Met	lyr	vai	Gly	GIY		GIU	1112	VI B	Leu		·
630					635		•			640	٠.			:	645	
	٠.				• •••	•	, ,		•					٠.		
gcc	gca	tgc	aac	ttc	act	cgt	ggg	gat	cgc	tgc	gac	ttg	gag	gac	agg	2323
Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Cýs	Asp	Leu	Glu	Asp	Arg	
				650					655					660		
gac	agg	agt	cag	ctg	tct	cct	ctg	ttg	cac	tct	acc	acg	gaa	tgg	gcc	2371
Asp	Arg	Ser	Gln	Leu	Ser	Pro	Leu	Leu	His	Ser	Thr	Thr	Glu	Trp	Ala	
_			665					670		•			675			
4			+ ~ ~		+00	+00	700	tta	000	act	++ a	tca	act	oot.	ctt	2419
TTe	Leu			Thr	lyr	Ser		Leu	Pro	Ата	Leu		1111	СТУ	Leu	
		680	1				685					690				
				•												
ctc	cac	ctt	cac	cag	aac	ato	gtg	gac	gta	caa	tac	atg	tat	ggc	ctc	2467
Leu	His	Leu	His	Gln	Asn	Ile	Val	Asp	Val	G1n	Tyr	Met	Tyr	· Gly	Leu	
	695					700)				705	;				
+00	00+	- mo+		905	, ,,,,	tac	gta	gtt	CDS	t.øc	i Gyb	tee	gts	g gta	ctc	2515
		Ala	1 TT6	e ini			. иал	. val	. vr.g			ıııþ	197	. val	Leu	
710	•				715	Ò				720)				725	

tta ttc ctg ctc tta gcg gac gcc aga gtc tgc gcc tgc ttg tgg atg Leu Phe Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp Met ctc atc ttg ttg ggc cag gcc gaa gca gca ttg gag aag ttg gtc gtc Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val ttg cac gct gcg agt gcg gct aac tgc cat ggc ctc cta tat ttt gcc Leu His Ala Ala Ser Ala Ala Asn Cys His Gly Leu Leu Tyr Phe Ala atc ttc ttc gtg gca gct tgg cac atc agg ggt cgg gtg gtc ccc ttg Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly Arg Val Val Pro Leu acc acc tat tgc ctc act ggc cta tgg ccc ttc tgc cta ctg ctc atg Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe Cys Leu Leu Met gca ctg ccc cgg cag gct tat gcc tat gac gca cct gtg cac gga cag Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala Pro Val His Gly Gln ata ggc gtg ggt ttg ttg ata ttg atc acc ctc ttc aca ctc acc ccg Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu Phe Thr Leu Thr Pro

ggg tat aag acc ctc ctc ggc cag tgt ctg tgg tgg ttg tgc tat ctc Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp Trp Leu Cys Tyr Leu ctg acc ctg ggg gaa gcc atg att cag gag tgg gta cca ccc atg cag Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp Val Pro Pro Met Gln gtg cgc ggc ggc cgc gat ggc atc gcg tgg gcc gtc act ata ttc tgc Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala Val Thr Ile Phe Cys ccg ggt gtg gtg ttt gac att acc aaa tgg ctt ttg gcg ttg ctt ggg Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu Leu Ala Leu Leu Gly cct gct tac ctc tta agg gcc gct ttg aca cat gtg ccg tac ttc gtc Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His Val Pro Tyr Phe Val aga gct cac gct ctg ata agg gta tgc gct ttg gtg aag cag ctc gcg Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu Val Lys Gln Leu Ala ggg ggt agg tat gtt cag gtg gcg cta ttg gcc ctt ggc agg tgg act Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala Leu Gly Arg Trp Thr ggc acc tac atc tat gac cac ctc aca cct atg tcg gac tgg gcc gct

Gly	Thr	Tyr	Ile	Tyr	Asp	His	Leu	Thr	Pro	Met	Ser	Asp	Trp	Ala	Ala	
950					955					960					965	
agc	ggc	ctg	cgc	gac	tta	gcg	gtc	gcc	gtg	gaa	ccc	atc	atc	ttc	agt	3283
Ser	G1y	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu	Pro	Ile	Ile	Phe	Ser	
				970					975					980		
			٠	· · ·	· · · ·	· .		•			•					
ccg	atġ	gag	aag	aag	gtc	atc	gtc	tgg	gga	gcg	gag	acg	gct	gca	tgt	3331
Pro	Met	Glu	Lys	Lys	Val	Ile	Val	Trp	Gly	Ala	Ġlu	Thr	Ala	Ala	Cys	*
٠		€.	985	٠.	•• ,	. ••	*	990		•	٠.	34.	995	••.	• '	
				•				٠.,	•							
ggg	gac	att	cta	cat	gga	ctt	ccc	gtg	tcc	gcc	cga	ctc	gġc	cag	gag	3379
Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala	Arg	Leu	Gly	G1n	Glu	
•		1000					1005					1010				
atc	ctc	ctc	ggc	cca	gct	gat	ggc	tac	acc	tcc	aag	ggg	tgg	aag	ctc	3427
Ile	Leu	Leu	Gly	Pro	Ala	Asp	Gly	Tyr	Thr	Ser	Lys	G1y	Trp	Lys	Leu	
	1015					1020		٠			1025					
ctt	gct	ccc	atc	act	gct	tat	gcc	cag	caa	aca	cga	ggc	ctc	ctg	ggc	3475
Leu	Ala	Pro	Ile	Thr	Ala	Tyr	Ala	Gln	Gln	Thr	Arg	Gly	Leu	Leu	Gly	
1030	0				1035	;				1040					1045	
gcc	ata	gtg	gtg	agt	atg	acg	ggg	cgt	gac	agg	aca	gaa	cag	gcc	ggg	3523
Ala	Ile	Val	. Val	Ser	Met	Thr	· Gly	Arg	Asp	Arg	Thr	G1u	G1n	Ala	Gly	
				1050)				1055					1060)	
																•
gaa	gto	caa	ato	cte	tcc	aca	gto	tct	cag	tcc	ttc	cto	gga	aca	acc	3571
													•		Thr	
•									/117					•		

PCT/JP2003/015038

atc tcg ggg gtt ttg tgg act gtt tac cac gga gct ggc aac aag act Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr

cta gcc ggc tta cgg ggt ccg gtc acg cag atg tac tcg agt gct gag Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu

ggg gac ttg gta ggc tgg ccc agc ccc cct ggg acc aag tct ttg gag Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu

ccg tgc aag tgt gga gcc gtc gac cta tat ctg gtc acg cgg aac gct Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala

gat gtc atc ccg gct cgg aga cgc ggg gac aag cgg gga gca ttg ctc Asp Val Ile Pro Ala Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu

tcc ccg aga ccc att tcg acc ttg aag ggg tcc tcg ggg ggg ccg gtg Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val

ctc tgc cct agg ggc cac gtc gtt ggg ctc ttc cga gca gct gtg tgc Leu Cys Pro Arg Gly His Val Val Gly Leu Phe Arg Ala Ala Val Cys

tct cgg ggc gtg gcc aaa tcc atc gat ttc atc ccc gtt gag aca ctc 3955 · Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu gac gtt gtt aca agg tct ccc act ttc agt gac aac agc acg cca ccg Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro gct gtg ccc cag acc tat cag gtc ggg tac ttg cat gct cca act ggc Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly agt gga aag agc acc aag gtc cct gtc gcg tat gcc gcc cag ggg tac Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr aaa gta cta gtg ctt aac ccc tcg gta gct gcc acc ctg ggg ttt ggg Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly gcg tac cta tcc aag gca cat ggc atc aat ccc aac att agg act gga Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg. Thr Gly gtc agg acc gtg atg acc ggg gag gcc atc acg tac tcc aca tat ggc Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr Tyr Ser Thr Tyr Gly

aaa	ttt	ctc	gcc	gat	gġg	ggc	tgc	gct	agc	ggc	gcc	tat	gac	atc	atc	4291
Lys :	Phe	Leu	Ala	Asp	Gly	Gly	Cys	Ala	Ser	Gly	Ala	Tyr	Asp	Ile	Ile	
		1	305				1	310				1	315			
ata	tgc	gat	gaa	tgc	cac	gct	gtg	gat	gct	acc	tcc	att	ctc	ggc	atc	4339
Ile	Cys	Asp	Glu	Cys	His	Ala	Val	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	
		1320				1	325					1330	F. +	•		
•					• • •							· · .	-	٠		
gga	acg	gtc	ctt	gat	caa	gca	gag	aca	gcc	ggg	gtc	aga	cta	act	gtg	4387
Gly	Thr	Val	Leu	Asp	Gl'n	Ala	G1u	Thr	Ala	G1y	Val	Arg	Leu	Thr	Val	
1	.335				;	1340				:	1345					
														٠		
ctg	gct	acg	gcc	aca	ccc	ccc	ggg	tca	gtg	aca	acc	ccc	cat	ccc	gat	4435
Leu	Ala	Thr	Ala	Thr	Pro	Pro	G1y	Ser	Val	Thr	Thr	Pro	His	Pro	Asp	
1350)				1355				:	1360					1365	
ata	gaa	gag	gta	ggc	ctc	ggg	cgg	gag	ggt	gag	atc	ccc	ttc	tat	ggg	4483
Ile	Glu	G1u	Val	Gly	Leu	Gly	Arg	Glu	Gly	G1u	Ile	Pro	Phe	туг	Gly	•
				1370					1375					1380)	
agg	gcg	att	ccc	cta	tcc	tgc	atc	aag	gga	ggg	aga	cac	cte	g ati	t ttc	4531
Arg	Ala	ıle	Pro	Leu	Ser	Cys	Ile	Lys	Gly	G1y	Arg	g His	: Lei	ı Ile	e Phe	
			1385	i				1390)				1395	5		
tgc	cac	tca	a aag	g aaa	aag	g tgt	gac	gag	ctc	gcg	gcg	g gco	c ct	t cg	g ggc	4579
Cys	His	s Sei	r Lys	Lys	s Lys	s Cys	Asp	Gl:	ı Leu	Ala	. Ala	a Ala	a Lei	u Ar	g Gly	
		1400)				1405	5				1410)			
atg	gg	c tt	g aat	t gco	c gt	g gca	a tac	tat	t aga	ggg	g tt	g ga	c gt	c tc	c ata	4627

Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Ile 1415 1420 1425

ata cca gct cag gga gat gtg gtg gtc gtc gcc acc gac gcc ctc atg

11e Pro Ala Gln Gly Asp Val Val Val Ala Thr Asp Ala Leu Met

1430

1445

acg ggg tac act gga gac ttt gac ccc gtg atc gac tgc aat gta gcg 4723

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala

1450 1455 1460

gtc acc caa gct gtc gac ttc agc ctg gac ccc acc ttc act ata acc 4771

Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr

1465 1470 1475

aca cag act gtc cca caa gac gct gtc tca cgc agt cag cgc cgc ggg 4819

Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly

1480 1485 1490

cgc aca ggt aga gga aga cag ggc act tat agg tat gtt tcc act ggt 4867

Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg Tyr Val Ser Thr Gly

1495 1500 1505

gaa cga gcc tca gga atg ttt gac agt gta gtg ctt tgt gag tgc tac 4915
Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr
1510 1515 1520 1525

gac gca ggg gct gcg tgg tac gat ctc aca cca gcg gag acc acc gtc 4963
Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro Ala Glu Thr Thr Val

PCT/JP2003/015038 WO 2004/104198

agg ctt aga gcg tat ttc aac acg ccc ggc cta ccc gtg tgt caa gac Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp cat ctt gaa ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp gcc cac ttc ctc tcc caa aca aag caa gcg ggg gag aac ttc gcg tac Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Glu Asn Phe Ala Tyr cta gta gcc tac caa gct acg gtg tgc gcc aga gcc aag gcc cct ccc Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro ccg tcc tgg gac gcc atg tgg aag tgc ctg gcc cga ctc aag cct acg Pro Ser Trp Asp Ala Met Trp Lys Cys Leu Ala Arg Leu Lys Pro Thr

ctt gcg ggc ccc aca cct ctc ctg tac cgt ttg ggc cct att acc aat Leu Ala Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Pro Ile Thr Asn

gag gtc acc ctc aca cac cct ggg acg aag tac atc gcc aca tgc atg Glu Val Thr Leu Thr His Pro Gly Thr Lys Tyr Ile Ala Thr Cys Met

caa g	gct	gac	ctt	gag	gtc	atg	acc	agc	acg	tgg	gtc	cta	gct	gga	gga	5347
Gln A	Ala .	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp	Val	Leu	Ala	Gly	Gly	
16	355				1	660				1	665					•
gtc	ctg	gca	gcc	gtc	gcc	gca	tat	tgc	ctg	gcg	act	gga	tgc	gtt	tcc	5395
Val 1	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Cys	Leu	Ala	Thr	Gly	Cys	Val	Ser	
1670	• •		. •	1	675	•	••••		.1	680	··. "·	•	•		1685	
•				i., -						•	÷ 🕶			2 - 2 * 1 2 - 2 * 1 2 - 2 * 1		
atc	atc	ggc	cgc	ttg	cac	gtc	aac	cag	cga	gtc	gtc	gtt	gcg	ccg	gat	5443
Ile	Ile	Gly	Arg	Leu	His	Val	Asn	Gln	Arg	Val	Val	Val	Ala	Pro	Asp	•
				1690				1	1695					1700		
•				•	•						•					
aag	gag	gtc	ctg	tat	gag	gct	ttt	gat	gag	atg	gag	gaa	tgc	gcc	tct	5491
Lys	Glu	Val	Leu	Tyr	G1u	Ala	Phe	Asp	Glu	Met	Glu	G1u	Cys	Ala	Ser	
			1705					1710					1715	•		
agg	gcg	gct	ctc	atc	gaa	gag	ggg	cag	cgg	ata	gcc	gag	atg	ttg	g aag	5539
Arg	Ala	Ala	Leu	Ile	G1u	G1u	Gly	Gln	Arg	Ile	Ala	Glu	Met	Leu	ı Lys	
		1720)				1725					1730)			
										•						
tcc	aag	ato	caa	ggc	ttg	ctg	cag	cag	gcc	tct	aag	g cag	ggc	cag	g gac	5587
Ser	Lys	$I1\epsilon$	e Gln	Gly	Leu	Leu	Gln	Gln	Ala	Ser	Lys	G1r	n Ala	a Glr	n Asp	
	1735					1740	ı				1745	5				
ata	caa	ccc	c gct	t atg	g cag	g gct	tca	tgg	ccc	aaa	a gtg	g gaa	a caa	a tt	t tgg	5635
Ile	G1n	Pro	o Ala	a Met	t Glr	ı Ala	Ser	Trp	Pro	Lys	s Val	l Glu	u Gli	n Ph	e Trp	
175	0				1755	5				1760)				1765	

	gcc	aga	cac	atg	tgg	aac	ttc	att	agc	ggc	atc	caa	tac	ctc	gca	gga	5683
	Ala	Arg	His	Met	Trp	Asn	Phe	Ile	Ser	Gly	Ile	Gln	Tyr	Leu	Ala	Gly	
1770										775				1	780		•
	ttg	tca	aca	ctg	cca	ggg	aac	ссс	gcg	gtg	gct	tcc	atg	atg	gċa	ttc	5731
	Leu	Ser	Thr	Ļeu	Pro	G1y	Asn	Pro	Ala	Val	Ala	Ser	Met	Met	Ala	Phe	
		~		1785		رم			1790					1795			
		***	٠	٥.			·e	٠ .					٠.		•	©	
,	agt	gCC	gcc	ctc	acc	agt	cčg	ttg	tcg	acc	agt	acc	acc	atc	ctt	ctc	5779
									Ser		٠.					• • •	•
			1800				Ť	1805	•				1810	.*			
													•				
	аас	atc	atg	gga	ggc	tgg	tta	gcg	tcc	cag	atc	gca	cca	ccc	gcg	ggg	5827
									Ser								
		1815 1815	2.2.2.2				1820					1825					
	·																
	gcc	acc	ggr	: t.t.t	gtc	gtc	agt	ggc	ctg	gtg	ggg	gct	gcc	gtg	ggo	agc	5875
																Ser	
	183		O.J.		,	1835					1840					1845	
	100	•						, .									
	ata	aac	cto	- oot	t aag	o ete	r cte	gte	gac	atc	ctg	gca	ı gga	ı tat	; ggt	t gcg	5923
																y Ala	
	116	GIY	Dec	1 01	1850		. дос	. ,	op	1855			•	•	1860		
					1000	,				1000							
		_ 4.4		~ ~~	~ ~~	o ete	a t		a ++c	. 220	r atr	. ate	t.c1	t ggo	e ga	g aag	5971
	Gly	116	e .Se:			a Let	ı va.	r Are		•	, TT	MIC		1879		u Lys	
				186	b				1870	,				1016	,		
															_ 1 -	+	6010
	ccc	tct	t at	g ga	a ga	t gt	c at	c aa	t cta	a ct	g cci	gg	g at	CCT	g tc	t ccg	6019

Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro 1880 1885 1890

gga gcc ctg gtg gtg ggg gtc atc tgc gcg gcc att ctg cgc cgc cac 6067 Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His 1895 1900 1905

gtz gga ccg ggg gag ggc gcg gtc caa tgg atg aac agg ctt att gcc 6115
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala
1910 1925

ttt gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag 6163
Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu
1930 1935 1940

tcg gat gcg tcg cag cgt gtg acc caa cta ctt ggc tct ctt act ata 6211 Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile 1945 1950 1955

acc agc cta ctc aga aga ctc cac aat tgg ata act gag gac tgc ccc 6259

Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro

1960 1965 1970

atc cca tgc tcc gga tcc tgg ctc cgc gac gtg tgg gac tgg gtt tgc 6307

Ile Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys

1975 1980 1985

acc atc ttg aca gac ttc aaa aat tgg ctg acc tct aaa ttg ttc ccc 6355

Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro

29/117

1990 1995 2000 2005

aag ctg ccc ggc ctc ccc ttc atc tct tgt caa aag ggg tac aag ggt 6403

Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly

2010 2015 2020

gtg tgg gcc ggc act ggc atc atg acc acg cgc tgc cct tgc ggc gcc 6451

Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala

2025 2030 2035

aac atc tct ggc aat gtc cgc ctg ggc tct atg agg atc aca ggg cct 6499

Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro

2040 2045 2050

aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgc tac 6547

Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr

2055 2060 2065

acg gag ggc cag tgc gcg ccg aaa ccc ccc acg aac tac aag acc gcc 6595

Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr Asn Tyr Lys Thr Ala

2070 2075 2080 2085

atc tgg agg gtg gcg gcc tcg gag tac gcg gag gtg acg cag cat ggg 6643

Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly

2090 2095 2100

tcg tac tcc tat gta aca gga ctg acc act gac aat ctg aaa att cct 6691
Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp Asn Leu Lys Ile Pro
2105 2110 2115

tgc caa cta cct tct cca gag ttt ttc tcc tgg gtg gac ggt gtg cag Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln atc cat agg ttt gca ccc aca cca aag ccg ttt ttc cgg gat gag gtc Ile His Arg Phe Ala Pro Thr Pro Les Pro Phe Phe Arg Asp Glu Val tcg ttc tgc gtt ggg ctt aat tcc tat gct gtc ggg tcc cag ctt ccc Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val Gly Ser Gln Leu Pro tgt gaa cct gag ccc gac gca gac gta ttg agg tcc atg cta aca gat Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg Ser Met Leu Thr Asp ccg ccc cac atc acg gcg gag act gcg gcg cgg cgc ttg gca cgg gga Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg Arg Leu Ala Arg Gly tca cct cca tct gag gcg agc tcc tca gtg agc cag cta tca gca ccg Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser Gln Leu Ser Ala Pro tcg ctg cgg gcc acc tgc acc acc cac agc aac acc tat gac gtg gac Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn Thr Tyr Asp Val Asp

atg	gtc	gat	gcc	aac	ctg	ctc	atg	gag	ggc	ggt	gtg	gct	cag	aca	gag	7075	
Met	Val	Asp	Ala	Asn	Leu	Leu	Met	Glu	Gly	Gly	Val	Ala	Gln	Thr	Glu		
2230)			2	235				2	240				2	2245		
				,													
cct	gag	tcc	agg	gtg	ccc	gtt	ctg	gac	ttt	ctc	gag	cca	atg	gcc	gag	7123	
Pro	Glu	Ser	Arg	Val	Pro	Va1	Leu	Asp	Phe	Leu	G1u	Pro	Met.	Ala	Glu		
			. 2	2250				. 2	2255	:			4	2260			
			•		-3°		· .	٠,							•		
gaa	gag	agc	gac	ctt	gag	cċc	tca	ata	cca	tcg	gag	tgc	atg	ctc	ccc	7171	
G1u	Glu	Ser	Asp	Leu	Glu	Pro	Ser	Ile	Pro	Ser	G1u	Cys	Met	Leu	Pro		
•			2265				. 2	2270					2275				
agg	agc	ggg	ttt	cca	cgg	gcc	tta	ccg	gct	tgg	gca	cgg	cct	gac	tac	7219	
Arg	Ser	G1y	Phe	Pro	Arg	Ala	Leu	Pro	Ala	Trp	Ala	Arg	Pro	Asp	Tyr		
		2280					2285					2290					
									•			•					
aac	ccg	ccg	ctc	gtg	gaa	tcg	tgg	agg	agg	cca	gat	tac	caa	ccg	ccc	7267	
Asn	Pro	Pro	Leu	Val	G1u	Ser	Trp	Arg	Arg	Pro	Asp	Tyr	Gln	Pro	Pro		
	2295					2300)				2305						
acc	gtt	gct	ggt	tgt	gct	cto	ccc	ccc	ccc	aag	aag	gcc	ccg	ace	g cct	7315	
Thr	Val	Ala	Gly	Cys	. Ala	Let	ı Pro	Pro	Pro	Lys	Lys	Ala	Pro	Thr	Pro		
2310 2						5				2320			2325				
ccc	cca	ı agg	g aga	ı cgo	cge	g aca	a gtg	ggt	ctg	ago	gag	ago	acc	ata	a tca	7363	
															e Ser		
				2330)				2335	5				2340)		
~ 00					a o t (m me.	2 2 2 4 (, acc	. ++1	- 000	. cas	7 000	a ccc	e teg	7411	

Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe Gly Gln Pro Pro Ser 2345 2350 2355

agc ggt gat gca ggc tcg tcc acg ggg gcg ggc gcc gcc gaa tcc ggc 7459

Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly Ala Ala Glu Ser Gly

2360 2365 2370

ggt ccg acg tcc cct ggt gag ccg gcc ccc tca gag aca ggt tcc gcc 7507 Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser Glu Thr Gly Ser Ala 2375 2380 2385

tcc tct atg ccc ccc ctc gag ggg gag cct gga gat ccg gac ctg gag 7555

Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu

2390 2395 2400 2405

tct gat cag gta gag ctt caa cct ccc ccc cag ggg ggg ggg gta gct 7603 Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Gly Val Ala 2410 2415 2420

ccc ggt tcg ggc tcg ggg tct tgg tct act tgc tcc gag gag gac gat 7651

Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp

2425 2430 2435

acc acc gtg tgc tgc tcc atg tca tac tcc tgg acc ggg gct cta ata 7699

Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile

2440 2445 2450

act ccc tgt agc ccc gaa gag gaa aag ttg cca atc aac cct ttg agt 7747 Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Pro Leu Ser

aac tcg ctg ttg cga tac cat aac aag gtg tac tgt aca aca tca aag Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys agc gcc tca cag agg gct aaa aag gta act ttt gac agg acg caa gtg Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe Asp Arg Thr Gln Val ctc gac gcc cat tat gac tca gtc tta aag gac atc aag cta gcg gct Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala tcc aag gtc agc gca agg ctc ctc acc ttg gag gag gcg tgc cag ttg Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu Glu Ala Cys Gln Leu act cca ccc cat tct gca aga tcc aag tat gga ttc ggg gcc aag gag Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly Phe Gly Ala Lys Glu gtc cgc agc ttg tcc ggg agg gcc gtt aac cac atc aag tcc gtg tgg Val Arg Ser Leu Ser Gly Arg Ala Val Asn His Ile Lys Ser Val Trp

gcc	aaa	aat	gag	gtg	ttc	tgc	gtg	gac	ccc	gcc	aag	ggg	ggt	aag	aaa	8131
Ala	Lys	Asn	Glu	Val	Phe	Cys	Val	Asp	Pro	Ala	Lys	Gly	Gly	Lys	Lys	
		2	585				2	590				2	595			
cca	gct	cgc	ctc	atc	gtt	tac	cct	gac	ctc	ggc	gtc	cgg.	gtc	tgc	gag	8179
Pro	Ala	Arg	Leu	Ile	Val	Tyr	Pro	Asp	Leu	Gly	Val	Årg	Val	Cys	Glu	
	2	2600	:			2	2605		·i,		, 2	2610				
					• .	, • • ·				· · · :	• :	.:		≱.		
aaa	atg	gcc	ctc	tat	gac	att	aca	caa	aag	ctt	cct	cag	gcg	gta	atg	8227
Lys	Met	Ala	Leu	Tyr	Asp	Ile	Thr	G1n	Lys	Leu	Pro	Gln	Ala	Val	Met	
2	2615				:	2620				2	2625					
gga	gct	tcc	tat	ggc	ttc	cag	tac	tcc	cct	gcc	caa	cgg	gtg	gag	tat	8275
Gly	Ala	Ser	Tyr	Gly	Phe	G1n	Tyr	Ser	Pro	Ala	G1n	Arg	Val	Glu	Tyr	
263	0				2635					2640					2645	
ctc	ttg	aaa	gca	tgg	gcg	gaa	aag	aag	gac	ccc	atg	ggt	ttt	tcg	tat	8323
Leu	Leu	Lys	Ala	Trp	Ala	Glu	Lys	Lys	Asp	Pro	Met	Gly	Phe	Ser	Tyr	
				2650					2655		•			2660		
gat	acc	cga	tgc	ttc	gac	tca	acc	gtc	act	gag	aga	gac	ato	agg	acc	8371
Asp	Thr	Arg	Cys	Phe	Asp	Ser	Thr	Val	Thr	Glu	Arg	, Asp	Ile	e Arg	Thr	
			2665	5				2670)				2675	5		
gag	gag	tcc	ata	ı tac	cag	g gcc	tgo	tcc	ctg	g ccc	gag	g gag	gco	cgc	act	8419
															g Thr	
		2680					2685					2690				

gcc	ata	cac.	tcg	ctg	act	gag	aga	ctt	tac	gta	gga	ggg	ccc	atg	ttc	8467
Aļa	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Tyr	Val	Gly	G1y	Pro	Met	Phe	
2	695				. 2	2700				2	705					
										•						
220	age	ឧឧଟ	øøt.	caa	acc.	tgc	ggt	tac	aġa	cgt	tgc	cgc	gcc	agc	ggg	8515
				Gln												
		LJS		-	2715				٠.	-		2			2725	
2710	,		•					•		31 2 V	₹¥ , * .	•		0		
		•	· ·	• •	•					•	•	, .		•		9569
									•						gcc	8563
Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Iḷe	Thr	Cys ·	Tyr	Val	Lys	Ala	
				2730				. 4	2735					2740		
cta	gcg	gcc	tgc	aag	gct	gcg	ggg	ata	gtt	gcg	ccc	aca	atg	ctg	gta	8611
Leu	Ala	Ala	Cys	Lys	Ala	Ala	G1y	Ile	Val	Ala	Pro	Thr	Met	Leu	Val	
			2745					2750					2755			
tac	aac	gat	. gac	cta	gta	gtc	atc	tca	gaa	agc	cag	ggg	act	gag	gag	8659
															ı Glu	
Cys	СТУ			, Dea	, , ,	141	2765		014	202		2770				
		2760	,				2100	l				2110				
														•	, ,	0707
															tct	8707
Asp	Glu	ı Arg	g Asr	ı Lev	ı Arg	; Ala	Phe	Thr	G1u	Ala	Met	Thr	r Arg	у Туз	Ser	
	2775	5				2780)				2785	5				
gcc	cct	t cc	t gg1	t gat	ccc	ccc	aga	а ссе	gaa	tat	gac	ct	g gag	g cta	a ata	8755
Ala	. Pro	o Pro	o G1:	y Asp) Pro	Pro	Arg	g Pro	Glu	ı Tyr	. Ası	Let	ı Glu	ı Lei	u Ile	
279					2799					2800					2805	
	. •															
	•		, .					.					a 0.00	~ ~~	0 000	8803
aca	a tc	c tg	t tc	c tca	a aat	t gt	g tc	ı gış	g gc	3 CT8	s gg	נטט נ	g cgi	5 88	c cgc	0003

Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Arg Gly Arg 2810 2815 2820

cgc aga tac tac ctg acc aga gac cca acc act cca ctc gcc cgg gct 8851

Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala

2825 2830 2835

gcc tgg gaa aca gtt aga cac tcc cct atc aat tca tgg ctg gga aac '8899
Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn Ser Trp Leu Gly Asn
2840 2845 2850

atc atc cag tat gct cca acc ata tgg gtt cgc atg gtc cta atg aca 8947

Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr

2855 2860 2865

cac ttc ttc tcc att ctc atg gtc caa gac acc ctg gac cag aac ctc 8995

His Phe Phe Ser Ile Leu Met Val Gln Asp Thr Leu Asp Gln Asn Leu

2870 2885 2880 2885

aac ttt gag atg tat gga tca gta tac tcc gtg aat cct ttg gac ctt 9043
Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Asn Pro Leu Asp Leu
2890 2895 2900

cca gcc ata att gag agg tta cac ggg ctt gac gcc ttt tct atg cac 9091
Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Met His
2905 2910 2915

aca tac tct cac cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa 9139
Thr Tyr Ser His His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys

2920

2925

2930

			0107
ctt ggg gcg cca ccc ctc agg gtg			9187
Leu Gly Ala Pro Pro Leu Arg Val	Trp Lys Ser	Arg Ala Arg Ala Val	
2935 2940	2	2945	
agg gcg tcc ctc atc tcc cgt gga	ggg aaa gcg	gcc gtt tgc ggc cga	923 5
Arg Ala Ser Leu Ile Ser Arg Gly	Gly Lys Ala	Ala Val Cys Gly Arg	
2950 2955	2960	2965	
			•
tat ctc ttc aat tgg gcg gtg aag	acc aag ctc	aaa ctc act cca ttg	9283
Tyr Leu Phe Asn Trp Ala Val Lys			
2970	2975	2980	
ccg gag gcg cgc cta ctg gac tta	tcc agt tgg	ttc acc gtc ggc gcc	9331
Pro Glu Ala Arg Leu Leu Asp Leu			
•	2990	2995	
2903			
200 200	ata tea eac	e goe ega ece ege tea	9379
ggc ggg ggc gac att ttt cac agc			•
Gly Gly Gly Asp Ile Phe His Ser			
3000 3005	1	3010	
			0497
tta ctc ttc ggc cta ctc cta ctt			9427
Leu Leu Phe Gly Leu Leu Leu Leu	Phe Val Gly	y Val Gly Leu Phe Leu	
3015 3020		3025	
ctc ccc gct cgg tag agcggcacac	actaggtaca	ctccatagct aactgttcct	9482
Leu Pro Ala Arg			

tcacggctag ctgtgaaagg tccgtgagcc gcatgactgc agagagtgcc gtaactggtc 9662

tctctgcaga tcatgt

96.78

<210> 4

<211> 3033

<212> PRT

<213> Hepatitis C virus

<400> 4

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

1 5 10 15

Arg Arg Pro Glu Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr
35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
50 55 60

Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly 65 70 75 80

Arg Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp

85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro

			100					105					110		
Arg	His	Arg	Ser	Arg	Asn	Val	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Cys
		115					120					125			
Gly	Phe	Ala	Asp	Leu	Met	G1 y	Tyr	Ile	Pro	Val	Val	Gly	Ala	Pro	Leu
	130					135				•	140				
Ser	Gly	Ala	Ala	Arg	Ala	Val	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp
145					150	•	· .	٠.		155		÷	. ·	•	160
Gly	Val	Asn	Tyr	Āla	Thr	Ĝly	Asn	Leu	$\dot{\tilde{P}}ro$	Gly	Phe	Pro	Phe	Ser	Ile
				165		•			170					175	
Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Île	Thr	Val	Pro	Val	Ser	Ala	Åla
			180					185					190		
Gln	Val	Lys	Asn	Thr	Ser	Ser	Ser	Tyr	Met	Val	Thr	Asn	Asp	Cys	Ser
		195					200					205			
Asn	Asp	Ser	Ile	Thr	Trp	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Val	Pro
	210					215					220				
Gly	Cys	Val	Pro	Cys	Glu	Arġ	Val	Gly	Asn	Thr	Ser	Arg	Cys	Trp	Val
225					230					235					240
Pro	Val	Ser	Pro	Asn	Met	Ala	Val	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Gln
				245					250					255	
Gly	Leu	Arg	Thr	His	Ile	Asp	Met	Val	Val	Met	Ser	Ala	Thr	Phe	Cys
			260					265					270		
Ser	Ala	Leu	Tyr	Val	Gly	Asp	Leu	Cys	Gly	G1y	Val	Met	Leu	Ala	Ala
		275					280					285		•	
Gln	Val	Phe	Ile	Val	Ser	Pro	G1n	Tyr	His	Trp	Phe	Val	Gln	Glu	Cys
	290					295					300				
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	Thr	Ile	Thr	Gly	His	Arg	Met	Ala	Trp
305					310					315					320
Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Ala	Thr	Met	Ile	Leu	Ala	Tyr
				325					330					335	

Val	Met	Arg	Val	Pro	Glu	Val	Ile	Ile	Asp	Ile	Val :	Ser	Gly	Ala	His
			340					345					350		
Trp	Gly	Val	Met	Phe	G1y	Leu	Ala	Tyr	Phe	Ser	Met	G1n	G1y	Ala	Trp
		355				•	360					365			
Ala	Lys	Val	Ile	Val	Ile	Leu	Leu	Leu	Ala	Ala	Gly ·	Val	Asp	Ala	Gly
	370					375					380		-	•	
Thr	Thr	Thr	Val	Gly	Gly	Ala	Val	Ala	Arg	Ser	Thr	Asn	Val	Ile	Ala
385	•			·	390	.				395	• •	. • • •			400
Gly	Val	Phe	Ser	His	Gly	Pro	G1n	Gln	Asn	Ile	G1n	Leu	Ile	Asn	Thr
· •			. :	405			• •	• •	410		. •	<		415	•
Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser
			420					425					430	·	
Leu	Asn	Thr	G1y	Phe	Leu	Ala	Ala	Leu	Phe	Tyr	Thr	Asn	Arg	Phe	Asn
		435					440					445			
Ser	Ser	G1y	Cys	Pro	Gly	Arg	Leu	Ser	Ala	Cys	Arg	Asn	Ile	Glu	Ala
	450					455					460				<u>.</u>
Phe	Arg	Ile	Gly	Trp	G1y	Thr	Leu	G1n	Tyr	Glu	Asp	Asn	Val	Thr	Asn
465					470					475				_	480
Pro	G1u	Asp	Met	Arg	Pro	Tyr	Cys	Trp			Pro	Pro	Lys		Cys
				485					490			_	_	495	
G1y	Val	. Val	Pro) Ala	Arg	Ser	· Val			Pro	Val	Tyr			Thr
			500					505					510		mı.
Pro	Ser	Pro	o Val	l Val	Val	Gly			Asp	Arg	Arg			Pro	Thr
		519					520			D 1		525		C	. TL
Tyr	Thi	r Tri	p G 1	y Glu	ı Asr			. Asp	Val	. Phe			ı Asn	ser ser	Thr
	530				_	535				m)	540		Λ	C	- Tha
Arg	g Pro	o Pro	o G1:	n Gly			o Ph€	e Gly	Cys			мет	. ASI	1 2e1	Thr
545					550				Б	555		. 1 01	. A	" A 1 <i>-</i>	560
G13	Ph	e Th	r Ly	s Th	r Cy:	s Gly	y Ala				s Arg	g ini	Are	g Ala	a Asp
								4	1/117	7					

				565					570					575	
Phe	Asn	Ala	Ser	Thr	Asp	Leu	Leu	Cys	Pro	Thr	Asp	Cys	Phe	Arg	Lys
			580				٠	585					590		
His	Pro	Asp	Ala	Thr	Tyr	Ile	Lys	Cys	Gly	Ser	G1y	Pro	Trp	Leu	Thr
		595					600			•		605			
Pro	Lys	Cys	Leu	Val	His	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Cys
	610	٠.		· .	•	615			•	•	620		•		:
Thr	Val	Asn	Phe	Thr	Ile	Phe	Lys	Ile	Arg	Met	Tyr	Val	Gly	Gly	Val
625	•		• • •		630					635			, . 		640
Ġlu	His	Arg	Leu	Thr	Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Árg	Cys
				645					650					655	
Asp	Leu	Glu	Asp	Arg	Asp	Arg	Ser	Gln	Leu	Ser	Pro	Leu	Leu	His	Ser
			660		·			665					670		•
Thr	Thr	Glu	Trp	Ala	Ile	Leu	Pro	Cys	Thr	Tyr	Ser	Asp	Leu	Pro	Ala
		675					680					685			
Leu	Ser	Thr	G1y	Leu	Leu	His	Leu	His	G1n	Asn	Ile	Val ·	Asp	Val	Gln
	690					695					700				
Tyr	Met	Tyr	G1y	Leu	Ser	Pro	Ala	Ile	Thr	Lys	Tyr	Va1	Val	Arg	Trp
705					710					715					720
G1u	Trp	Val	Val	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg		Cys
				725					730					735	
Ala	Cys	Leu	Trp	Met	Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu			Leu
			740					745					750		
Glu	Lys	Leu	ı Val	. Val	Leu	His	. Ala	Ala	Ser	· Ala	Ala			His	Gly
		755					760					765			
Leu	Let	і Туг	Phe	e Ala	lle	Ph€	Phe	Val	Ala	ı Ala			s Ile	Arg	Gly
	770					775					780			_	
Arg	y Val	l Val	l Pro	Leu	ı Thr	Thi	тук	Cys	Let			7 Leu	ı Trp	Pro	Phe
785	5				790)				795	5				800

_	T	T	T	M	47 -	T	Dage	Λ	C1 5	۸1a	Tur	412	Tur	Acn	ٔ د ۲ ۵
Cys	Leu	Leu	Leu		Ala	Leu	Pro			MIA	1) 1	nia	1) 1		nra
				805					810			_		815	_
Pro	Val	His	Gly	Gln	Ile	Gly	Val	G1y	Leu	Leu	Ile	Leu	Ile	Thr	Leu
			820					825					830		
Phe	Thr	Leu	Thr	Pro	Gly	Tyr	Lys	Thr	Leu	Leu	Gly	Gln	Cys	Leu	Trp
		835	•				840	. • •				845			
Trp	Leu	Cys	Tyr	Leu	Leu	Thr	Leu	Gly	Glu	Ala	Met	Ile	Gln	Glu	Trp
٠.	850	•	•			855			٠		860	•	. *.	•.	
Val	Pro	Pro	Met	Gln	Val	Årg	Gly	G1y	Arg	Asp	Gly	Ile	Ala	Trp	Ala
865	:		٠.٠	•	870			. • .	·	875		٠	. :	٠	880
Val	Thr	Ile	Phe	Cys	Pro	Gly	Val	Val	Phe	Asp	Ile	Thr	Lys	Trp	Leu
				885					890					895	
Leu	Ala	Leu	Leu	G1y	Pro	Ala	Tyr	Leu	Leu	Arg	Ala	Ala	Leu	Thr	His
			900					905					910		
Val	Pro	Tyr	Phe	Val	Arg	Ala	His	Ala	Leu	Ile	Arg	Val	Cys	Ala	Leu
		915					920					925			
Val	Lvs			Ala	G1v	G1 y	Arg	Tyr	Val	Gln	Val	Ala	Leu	Leu	Ala
, 42	930					935					940		-		
Lau			Trr	Thr	GT v		· Tyr	Ile	Tvr	Asp	His	Leu	Thr	Pro	Met
945		111 6	, 111	, 1111	950		-,-		-,-	955					960
		ጥ ъъ	. 41.	. 110			. Leu	Aro	Asn			Va1	Ala	Val	
Ser	кър	111) Ala			GIY	Leu		970			, , , ,		975	
	*1	*1	D1	965		Mad	. (1	Lva			Tla	. Val	Trn		
Pro) IIe	116			rrc	ме (: Glu			val	. 116	, va.	990		nia
			980				T 1	985		0.1	•	. D			. 41.
Glu	ı Thr			a Cys	s Gly	, Ası	o Ile		HIS	з СТУ	Let			Set	Ala
		998					1000					1005			
Arg	g Leu	ı Gl	y ∙G1ı	n Glu	ı Ile	e Le	ı Let	ı Gly	Pro) Ala			y Tyr	Thi	Ser
	1010					101					1020				
Lys	s Gly	Tr	D Ly	s Lei	ı Let	ı Ala	a Pro) Ile	e Thr	c Ala	а Туз	r Ala	a Glr	n Glr	1 Thr

1025				1	030				1	035				1	040
Arg (Gly I	Leu I	Leu	G1y	Ala	Ile	Val	Val	Ser	Met	Thr	Gly	Arg	Asp	Arg
			1	045				1	050				1	055	
Thr (Glu (Gln A	Ala	Gly	Glu	Val	Gln	Ile	Leu	Ser	Thr	Val	Ser	G1n	Ser
		10	060				1	065				1	070		
Phe l	Leu	Gly '	Thr	Thr	Ile	Ser	Gly	Val	Leu	Trp:	Thr	Val	Tyr	His	Gly
	1	075		•		. 1	1080	÷	•	•	1	1085	•		
Ala	Gly	Asn	Lys	Thr	Leu	Ala	G1 y	Leu	Arg	Gly	Pro	Val	Thr	G1n	Met
İ	090	•		· .·	. 1	.095					1100			·· .	
Tyr	Ser,	Ser	Ala	G1u	Gly	Asp	Leu	Val	Gly	Trp	Pro	Ser	Pro	Pro	Gly
1105				1	1110				:	1115]	1120
Thr	Lys	Ser	Leu	Glu	Pro	Cys	Lys	Cys	Gly	Ala	Val	Asp	Leu	Tyr	Leu
]	1125]	130				-	1135	
Val	Thr	Arg	Asn	Ala	Asp	Val	Ile	Pro	Ala	Arg	Arg	Arg	Gly	Asp	Lys
		1	140					1145					1150		
Arg	Gly	Ala	Leu	Leu	Ser	Pro	Arg	Pro	Ile	Ser	Thr	Leu	Lys	Gly	Ser
	1	1155					1160					1165			
Ser	G1y	·Gly	Pro	Val	Leu	Cys	Pro	Arg	Gly	His	Val	Val	Gly	Leu	Phe
1	170					1175					1180				
Arg	Ala	Ala	Val	Cys	Ser	Arg	Gly	Val	Ala	Lys	Ser	Ile	Asp	Phe	Ile
1185	5				1190					1195	•				1200
Pro	Val	Glu	Thr	Leu	Asp	Val	Val	Thr	Arg	Ser	Pro	Thr	Phe	Ser	Asp
				1205				•	1210)				1215	;
Asn	Ser	Thr	Pro	Pro	Ala	. Val	Pro	Gln	Thr	Tyr	Glr.	l Val	Gly	Tyr	Leu
			1220	•				1225	;				1230	1	
His	Ala	Pro	Thr	Gly	Ser	Gly	/ Lys	Ser	Thr	Lys	· Val	Pro	Val	. Ala	Tyr
		1235					1240)				1245	5		
Ala	Ala	Gln	Gly	Туг	Lys	Val	l Leu	ı Val	. Leı	ı Asr	n Pro	Ser	· Val	. Ala	a Ala
	1250					125	5				1260)			

Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Leu	Ser	Lys	Ala	His	Gly	Ile	Asn	Pro
1265				1	270				1	275				1	280
Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Val	Met	Thr	G1y	Glu	Ala	Ile	Thr
]	1285				1	290				1	295	
Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	Gly	G1 _. y	Cys	Ala	Ser	Gly
		. 1	300				1	305				. 1	310		
Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ala	Val	Asp	Ala	Thr
		1315		· ·	•		1320					1325			
Ser	Ile	Leu	G1 _. y	Ile	Gly	Thr	Val	Leu	Asp	G1n	Ala	Glu	Thr	Ala	Gly
. ¨` İ	330	•:	•			1335			. •		1340	•	-		٠.
Val	Arg	Leu	Thr	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr
1345	5			:	1350				:	1355]	L360
Thr	Pro	His	Pro	Asp	Ile	Glu	Glu	Val	Gly	Leu	G1y	Arg	Glu	Gly	Glu
				1365]	1370				:	1375	
Ile	Pro	Phe	Tyr	Gly	Arg	Ala	Ile	Pro	Leu	Ser	Cys			Gly	Gly
			1380					1385					1390		
Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys			Glu	Leu	Ala
		1395					1400					1405		•	
Ala	Ala	Leu	Arg	Gly	Met	Gly	Leu	Asn	Ala				Tyr	Arg	Gly
	1410					1415					1420				
Leu	Asp	Val	Ser	· Ile	Ile	Pro	Ala	Gln	Gly			Val	Val		
142					1430					1435			_		1440
Thr	Asp	Ala	Leu	ı Met	Thr	Gly	7 Tyr				Phe	Asp			
				1445					1450					1455	
Asp	Cys	Asn	\Val	l Ala	l Val	l Thi	r Gln			. Asp	Phe	Ser			Pro
			1460					1465					1470		_
Thr	Phe	Thr	: I1e	e Thr	Thi	r Gli	n Thr	· Val	Pro	Glr.	ı Asp			Ser	Arg
		1475				•						1485			
Ser	Glr	n Arg	g Ar	g Gly	7 Ar	g Thi	r Gly	7 Arg	G13	7 Arg	g Glr	ı Gly	Thr	Tyr	Arg
								45	5/117	7					

1	490	•			1	495				1	500				
Tyr	Val	Ser	Thr	Gly	Glu	Arg	Ala	Ser	Gly	Met	Phe	Asp	Ser	Val	Val
1505				1	510				1	515				1	520
Leu	Cys	G1u	Cys	Tyr	Asp	Ala	Gly	Ala	Ala	Trp	Tyr	Asp	Leu	Thr	Pro
			1	525				1	530				1	535	
Ala	G1u	Thr	Thr	Va1	Arg	Leu	Arg	Ala	Tyr	Phe	Asn	Thr	Pro	Gly	Leu
		1	540	. `	•		. 1	545		•	•		1550		
Pro	Val	Cys	G1n	Asp	His	Leu	G1u	Phe	Trp	Glu	Ala	۷a1	Phe	Thr	Gly
٠.	1	555	• •			. 1	560		*			1565			
Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	Ġ1n	Thr	Lys	G1n	Ala	Gly
1	570					1575	•			:	1580				
Glu	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Tyr	G1n	Ala	Thr	Val	Cys	Ala	Arg
1585	5				1590				:	1595]	1600
Ala	Lys	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Ala	Met	Trp	Lys	Cys	Leu	Ala
				1605]	1610				,	1615	
Arg	Leu	Lys	Pro	Thr	Leu	Ala	Gly	Pro	Thr	Pro	Leu	Leu	Tyr	Arg	Leu
			1620			,		1625					1630		
G1y	Pro	Ile	Thr	Asn	Glu	Val	Thr	Leu	Thr	His	Pro	Gly	Thr	Lys	Tyr
	Ċ	1635					1640					1645			
Ile	Ala	Thr	Cys	Met	G1n	Ala	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp
	1650					1655					1660	ı			
Val	Leu	Ala	Gly	Gly	Val	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Cys	Leu	Ala
166	5				1670	ı				1675	i				1680
Thr	G1y	Cys	Val	Ser	Ile	Ile	Gly	Arg	Leu	His	Val	Asn	Gln	Arg	Val
				1685					1690)				1695	
Val	Val	Ala	Pro	Asp	Lys	Glu	Val	Leu	Tyr	Glu	ı Ala	Phe	e Asp	Glu	Met
			1700)				1705					1710)	
Glu	G1u	Cys	s Ala	a Ser	Arg	, Ala	·Ala	Leu	Ile	Glu	ı Glu	ı Gly	Gln	Arg	Ile
		1715	5				1720)				1729	5		

Ala G	lu Met	Leu	Lys	Ser I	.ys 1	[le (Gln G	ly l	Leu 1	Leu	Gln	Gln A	Ala S	er
173	30			17	735				1	740				
Lys G	ln Ala	Gln	Asp	Ile (Gln F	Pro	Ala M	let (Gln .	Ala	Ser	Trp	Pro L	.ys
1745			1	750				1	755				17	760
Val G	lu Gln	Phe	Trp	Ala	Arg I	His	Met 1	rp .	Asn	Phe	Ile	Ser	Gly I	[le
	*		1765	:	. •		17	770				1	775	
Gln T	yr Leu	Ala	Gly	Leu	Ser	Thr	Leu I	Pro	Ģly	Asn	Pro	Ala	Val A	Ala
		1780	٠			1	785		•	-		1790		
Ser M	let Met	Ala	Phe	Ser	Ala	Ala	Leu '	Thr	Ser	Pro	Leu	Ser	Thr S	Ser
	1795	;	٠		1	800		,	٠.	•	1805			•
Thr T	Thr Ile	Leu	Leu	Asn.	Ile	Met	Gly (Gly	Trp	Leu	Ala	Ser	Gln	Ile
· 18	310			1	815					1820				
Ala F	Pro Pro	Ala	. Gly	Ala	Thr	G1y	Phe	Val	Val	Ser	Gly	Leu	Val	Gly
1825				1830]	1835				1	840
Ala A	Ala Va	l G13	/ Ser	Ile	Gly	Leu	Gly	Lys	Val	Leu	Val	Asp	I1e	Leu
			1845				1	.850				į	1855	
Ala	Gly Ty	r Gly	y Ala	G1y	Ile	Ser	Gly	Ala	Leu	Val	Ala	Phe	Lys	Ile
		1860	0				1865	•				1870		
Met	Ser Gl	y G1	u Lys	Pro	Ser	Met	Glu	Asp	Val	Ile	Asr	ı Leu	Leu	Pro
	187	5				1880	١				1885	5	•	
Gly	Ile Le	u Se	r Pro	Gly	Ala	Leu	Val	Val	G1y	Val	. I1e	e Cys	Ala	Ala
1	890				1895					1900)			
Ile	Leu Ar	g Ar	g His	s Val	G1y	Pro	Gly	Glu	Gly	Ala	a Va	l _. G1n	Trp	Met
1905	5			1910	ı				1915	5			:	1920
Asn	Arg Le	eu Il	e Ala	a Phe	Ala	Ser	Arg	G1y	Asr	ı His	s Va	l Ala	Pro	Thr
			1929	5				1930)				1935	
His	Tyr Va	al Th	ır Glı	u Ser	Asp	A1a	a Ser	G1r	n Arg	g Va	1 Th	r Glr	ı Leu	Leu
		194					1945					1950		
Glv														~ -
OLY	Ser L	eu Th	ır Il	e Thr	Ser	: Le	u Leu	Arg	g Ar	g Le	u Hi	s Ası	ı Trp	TTe

	Ċ	1955					1960				:	1965			
Thr	Glu	Asp	Cys	Pro	Ile	Pro	Cys	Ser	Gly	Ser	Trp	Leu	Arg	Asp	Val
	1970					1975					1980				
Trp	Asp	Trp	Val	Cys	Thr	Ile	Leu	Thr	Asp	Phe	Lys	Asn	Trp	Leu	Thr
198	5			:	1990				:	1995			•	2	2000
Ser	Lys	Leu	Phe	Pro	Lys	Leu	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Cys	G1n
		٠		2005				. 4	2010		•			2015	•
Lys	Gly	Tyr	Lys	Gly	Val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg
			2020		~ .	. •		2025					2030		
Cys	Pro	Cys	Gly	Ala	Asn	Ile	Ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met
	4	2035				2	2040			•	2	2045			
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Cys	Met	Asn	Thr	Trp	Gln	G1y	Thr	Phe
2	2050				4	2055				2	2060				
Pro	Ile	Asn	Cys	Tyr	Thr	Glu	Gly	Gln	Cys	Ala	Pro	Lys	Pro	Pro	Thr
2065	5			2	2070				2	2075				2	2080
Asn	Tyr	Lys	Thr	Ala	Ile	Trp	Arg	Val	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
			2	2085				4	2090				2	2095	
Val	Thr	Gln	His	Gly	Ser	Tyr	Ser	Tyr	Val	Thr	Gly	Leu	Thr	Thr	Asp
		2	2100				2	2105				2	2110		
Asn	Leu	Lys	Ile	Pro	Cys	G1n	Leu	Pro	Ser	Pro	G1u	Phe	Phe	Ser	Trp
	2	2115				2	2120				2	2125			
Va1	Asp	Gly	Val	Gln	Ile	His	Arg	Phe	Ala	Pro	Thr	Pro	Lys	Pro	Phe
2	2130				2	2135				2	2140				
Phe	Arg	Asp	Glu	Val	Ser	Phe	Cys	Val	G1y	Leu	Asn	Ser	Tyr	Ala	Val
2145	5			2	2150				2	2155				2	2160
Gly	Ser	G1n	Leu	Pro	Cys	Glu	Pro	Glu	Pro	Asp	Ala	Asp	Val	Leu	Arg
			2	2165				2	2170				2	2175	
Ser	Met	Leu	Thr	Asp	Pro	Pro	His	Ile	Thr	Ala	Glu	Thr	Ala	Ala	Arg
		2	2180				2	2185				2	2190		

Arg l	Leu	Ala	Arg	Gly 3	Ser	Pro	Pro	Ser	Glu	Ala	Ser	Ser	Ser	Val	Ser
	2	195				2	200				2	205			
Gln]	Leu	Ser	Ala	Pro	Ser	Leu	Arg	Ala	Thr	Cys	Thr	Thr	His	Ser	Asn
2	210				2	215				2	220				
Thr '	Tyr	Asp	Val	Asp	Met	Val	Asp	Ala	Asn	Leu	Leu	Met	Glu	Gly	Gly
2225		٠		2	230				2	2235				2	240
Val	Ála	Gĺn	Thr	G1u	Pro	Glu	Ser	Årg	Val	Pro	Val	Leu-	Asp	Phe	Leu
		* .		2245				. 2	2250		•		2	2255	
Glu .	Pro	Met	Ala	G1u	Glu	G1u	Ser	Asp	Leu	Glu	Pro	Ser	Ile	Pro	Ser
	:	2	2260	.).			2	2265		••		. 2	270		
Glu	Cys	Met	Leu	Pro	Arg	Ser	Gly	Phe	Pro	Arg	Ala	Leu	Pro	Ala	Trp
	2	2275				2	2280	٠			2	2285			
Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Arg	Arg	Pro
2	290				2	2295	•			2	2300				
Asp	Tyr	Gln	Pro	Pro	Thr	Val	Ala	Gly	Cys	Ala	Leu	Pro	Pro	Pro	Lys
2305	5			2	2310				4	2315				2	2320
Lys	Ala	Pro	Thr	Pro	Pro	Pro	Arg	Arg	Arg	Arg	Thr	Val	G1y	Leu	Ser
				2325					2330					2335	
Glu	Ser	Thr	Ile	Ser	Glu	Ala	Leu	G1n	G1n	Leu	Ala	Ile	Lys	Thr	Phe
			2340)				2345					2350		
Gly	G1n	Pro	Pro	Ser	Ser	G1y	Asp	Ala	Gly	Ser	Ser	Thr	G1y	Ala	Gly
		2355	i				2360					2365			
Ala	Ala	Glu	Ser	Gly	G1 y	Pro	Thr	Ser	Pro	Gly	Glu	Pro	Ala	Pro	Ser
2	2370)				2375					2380)			
Glu	Thr	Gly	Sei	. Ala	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	Pro	Gly
238	5				2390)				2395	;				2400
Asp	Pro	Asp	Lei	ı Glu	Ser	: Asp	Glr	ı Val	. G1:	ı Lev	Gln	Pro	Pro	Pro	Gln
				2405	;				2410)				2415	
Gly	Gly	7 Gl3	/ Vai	l Ala	Pro	Gly	seı	Gl ₃	r Sei	r Gly	7 Ser	Trp	Ser	Thr	Cys
								49	9/117	7					

		24	420			•	2	425				2	430		
Ser (Glu	Glu A	Asp	Asp '	Thr T	hr V	/al	Cys (Cys	Ser l	Met	Ser	Tyr	Ser'	Trp
	2	435				24	440				2	445			
Thr	Glý	Ala	Leu	Ile	Thr F	ro (Cys	Ser 1	Pro	Glu (Glu	Glu	Lys	Leu	Pro-
2	450				24	155				2	4 60				
Ile	Asn	Pro	Leu	Ser	Asn S	Ser :	Leu	Leu .	Arg	Tyr	His	Asn	Lys.	Val	Tyr
2465				2	470		•	•	2	2475		• •	٠.	. Ż	480
Cys	Thr	Thr	Ser	Lys	Ser	Ala	Ser	Gln	Arg	Ala	Lys	Lys	Val	Thr	Phe
	٠		2	2485			•	2	490	• .			2	495	
Asp	Arg	Thr	G1n	Val	Leu	Asp	Ala	His	Tyr	Asp	Ser	Val	Leu	Lys	Asp
		2	2500				2	2505				2	2510		
Ile	Lys	Leu	Ala	Ala	Ser	Lys	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu
		2515				2	2520				. :	2525			
Glu	Ala	Cys	Gln	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	Ser	Lys	Tyr _.	Gly
2	2530				2	535				2	2540				
Phe	Gly	Ala	Lys	Glu	Val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His
254	5			:	2550					2555					2560
Ile	Lys	Ser	Val	Trp	Lys	Asp	Leu	Leu	Glu	Asp	Pro	G1n	Thr	Pro	Ile
				2565				:	2570)				2575	
Pro	Thr	Thr	Ile	Met	Ala	Lys	Asn	Glu	Val	Phe	Cys	Val	Asp	Pro	Ala
			2580)				2585					2590		
Lys	G1 y	Gly	Lys	Lys	Pro	Ala	Arg	Leu	Ιlϵ	Val	Tyr	Pro) Asp	Leu	Gly
		2595					2600)				2605	5		
Val	Are	y Val	Cys	s Glu	Lys	Met	Ala	Leu	Туг	Asp	Ile	Thr	Gln	Lys	Leu
	2610)				2615					2620)			
Pro	Glr	n Ala	\Va	l Met	Gly	Ala	Ser	Tyr	Gly	y Phe	Glr	туз	r Ser	Pro	Ala
262	5				2630					2635	5				2640
G1n	Ar	g Val	l Gl	u Tyr	Leu	Leu	ı Lys	s Ala	Tr	p Ala	Glu	ı Ly:	s Lys	s Asp	Pro
				2645	5				265	0				2655	5

Met	Gly	Phe	Ser	Tyr	Asp	Thr	Arg	Cys	Phe	Asp	Ser	Thr	Val	Thr	G1u
		2	660				2	665				2	670		
Arg	Asp	Ile	Arg	Thr	Glu	Glu	Ser	Ile	Tyr	G1n	Ala	Cys	Ser	Leu	Pro
	2	2675				2	2680				2	2685			
G1u	G1u	Ala	Arg	Thr	Ala	Ile	His	Ser	Leu	Thr	G1u	Arg	Leu	Tyr	Val
. 2	2690	•			2	2695		•		2	2700				
Gly	Gly	Pro	Met	Phe	Asn	Ser	Lys	Gly	Gln	Thr	Cys	Gly	Tyr	Arg	Àrg
270	5	· .	-	2	2710		•		, 4	2715				2	2720
Cys	Arg	Ala	Ser	Gly	Val	Leu	Thr	Thr	Ser	Met	G1y	Asn	Thr	Ile	Thr
	-	*		2725	-				2730		٠.		. 4	2735	
Cys	Tyr	Val	Lys	Ala	Leu	Ala	Ala	Cys	Lys	Ala	Ala	Gly	Ile	Val	Ala
		:	2740				4	2745				2	2750	•	
Pro	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ile	Ser	G1u	Ser
	:	2755					2760				:	2765			
Gln	-Gly	Thr	G1u	Glu	Asp	G1u	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
	2770					2775					2780				
Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	G1u	Tyr
278	5				2790					2795					2800
Asp	Leu	Glu	Leu	Ile	Thr	Ser	Cys	Ser	Ser	Asn	Val	Ser	Val	Ala	Leu
				2805					2810	ı				2815	
G1y	Pro	Arg	Gly	Arg	Arg	Arg	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr
			2820)				2825					2830		
Pro	Leu	Ala	Arg	Ala	Ala	Trp	G1u	Thr	Val	Arg	His	Ser	Pro	Ile	Asn
		2835	i				2840	•	•			2845			
Ser	Trp	Leu	Gly	Asn	Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg
	2850)				2855					2860	•			
Met	: Val	Leu	Met	Thr	His	Phe	Phe	Ser	· Ile	e Leu	ı Met	. Val	Gln	Asp	Thr
286	6 5				2870)				2875	;				2880
Leu	ı Asp	Glr	ı Asr	ı Let	ı Asr	Phe	e Glu	ı Met	Туг	c G1y	Ser	· Val	Tyr	Ser	· Val
								51	l/117	7					

2885 2890 2895

Asn Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp 2900 2905 2910

Ala Phe Ser Met His Thr Tyr Ser His His Glu Leu Thr Arg Val Ala 2915 2920 2925

Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser 2930 2935 2940

Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala 2945 2950 2955 2960

Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu 2965 2970 2975

Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp
2980 2985 2990

Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg 2995 3000 3005

Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Phe Val Gly 3010 3015 3020

Val Gly Leu Phe Leu Leu Pro Ala Arg 3025 3030

<210> 5

<211> 9674

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341).. (9442)

<40	0> 5															
acc	cgcc	cct	aata	gggg	cg a	cact	ccgc	c at	gaat	cact	ccc	ctgt	gag	gaac	tactgt	60
ctt	cacg	cag	aaag	cgtc	ta g	ccat	ggcg	t ta	gtat	gagt	gtc	gtac	agc	ctcc	aggccc	120
				•												
ccc	cctc	ccg	ggag	agcc	at a	gtgg	tctg	c gg	aacc	ggtg	agt	acac	cgg	aatt	gccggg	180
	00±~		'	+ ~ + + .					, 44	,						0.40
aag	ac i g	RR Γ΄						•		cccg	gcc	attt.	ggg	cgtg	ccccg	240
caa	gact	gct					ggtt:			gcct	tgt	ggta	ctg	cctg	ataggg	300
				•		•		·			•					
tgc	ttgc	gag	tgcc	ccgg	ga g	gtct	cgtag	g ac	cgtg	cacc	atg	agc	aca	aat	ccc	355
											Met	Ser	Thr	Asn	Pro	
											1				. 5	
aaa	cct	caa	aga	aaa	acc	aaa	aga	aac	act	aac	cgt	cgc	cca	caa	gac	403
														Gln		
				10					15					20		
gtt	aag	ttt	ccg	ggc	ggc	ggc	cag	atc	gtt	ggc	gga	gta	tac	ttg	ttg	451
Val	Lys	Phe	Pro	Gly	Gly	G1y	Gln	Ile	Val	Gly	Gly	Val	Tyr	Leu	Leu	
			25					30					35			
CCg	CgC	agg	ggc	cee	agg	t.t.ø	ggt	σtσ	cgc	oco	aca	agg	220	gct	tea	499
														Ala		100

gag cgg tcc cag cca cgt ggg agg cgc cag ccc atc ccc aaa cat cgg 547 53/117

50

45

40

Glu	Arg	Ser	Gln	Pro	Arg	G1y	Arg	Arg	G1n	Pro	Ile	Pro	Lys	His .	Arg	
	55					60					65					
											•					
cgc	tcc	act	ggc	aag	tcc	tgg	ggg	aag	cca	gga	tac	ccc	tgg	ccc	ctg	595
Arg	Ser	Thr	G1y	Lys	Ser	Trp	G1y	Lys	Pro	Gly	Tyr	Pro	Trp	Pro	Leu	
70					75					80					85	
							•		٠.,		٠.	~ е		•		•
tat	ggg	aat	gag	ggg	ctc	ggt	tgg	gca	gga	tgg	ctc	cṫg	tcc	cct	cga	643
														Pro		
	, All		•	90						••	. ^t	•	•	100	: :	
		٠											•	·		•
ggt	tcc	cgt	ccc	tca	tgg	ggc	ccc	aat	gac	ccc	cgg	cat	agg	tcg	cgc	691
														Ser		
3			105		-			110					115			
																•
aat	gtg	ggt.	aag	gtc	atc	gat	acc	cta	acg	tgc	ggc	ttt	gcc	gac	ctc	739
														Asp		
	, 42	120					125					130				
††ø	000	tac	gtc	ccc	gtc	gta	ggc	gcc	CCS	ctt	agt	ggc	gtt	gcc	agt	787
														Ala		
Leu	135		, 41	110		140					145					
	100		٠			140							•			
	. 1 .				~ 1 ~	0.00	~+ ^	o t a			aaa	- σ+ +	aat	ttt	gca	835
														ttt Phe		
		ı Ala	i His	: СТУ			vai	Leu	GIL			Val	. ASI	. Phe	165	
150	•				155					160					100	
									,		11				. 4 -	009
														gcc		883
Thr	· G1v	7 Asr	ı Leu	ı Pro	Gly	Cys	: Ser	: Phe	e Sei	: ITe	Phe	e Leu	ı Lei	ı Ala	Leu	

170 175 180

ctg tcc tgc atc act act ccg gtc tct gct gtc caa gtg aag aac acc 931 Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val Gln Val Lys Asn Thr 185 190 195

agc aac gcc tat atg gcg act aac gac tgt tcc aat gac agc atc act 979
Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr

200 205 210

tgg cag ctt gag gcc gca gtc ctc cat gtc ccc ggg tgc gtc ccg tgc 1027.

Trp Gln Leu Glu Ala Ala Val Leu His Val Pro Gly Cys Val Pro Cys
215 220 225

gag aaa atg ggg aac aca tca cgg tgc tgg ata cca gtc tca cca aac 1075 Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile Pro Val Ser Pro Asn 230 245

gtg gct gtg cgg cag cct ggc gcc ctc acg cgg ggc ttg cgg acg cac 1123
Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg Gly Leu Arg Thr His
250 255 260

atc gac atg gtc gtg ttg tcc gcc acg ctc tgc tcc gct ctc tac gtg 1171

Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys Ser Ala Leu Tyr Val

265 270 275

ggg gac ctc tgt ggc ggg gtg atg ctc gcg tcc cag atg ttc att gtc 1219 Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser Gln Met Phe Ile Val 280 285 290

tcg	ccg	cag	cac	cac	tgg	ttc	gtg	cag	gaa	tgc	aat	tgc	tcc	atc	tac	1267
Ser 1	Pro	G1n	His	His	Trp	Phe	Val	Gln	Glu	Cys	Asn	Cys	Ser	Ile	Tyr	
	295				•	300					305					
cct	ggç	gcc	atc	act	ggg	cac	cgt	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	G1y	Ala	lle	Thr	Gly	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310			•		315	•				320					325	
•						•						. · · · ·		• •		
tgg.	tcg	CCC	acg	acc	acc	atg	atc	ctg	gcg	tác	gtg	atg	cgc	gtt	ccc	1363
						Met										
				330					335					340		
gag	gtc	atc	ata	gac	atc	att	agc	gga	gct	cac	tgg	ggc	gtc	atg	ttt	1411
Glu	Va1	Ile	Ile	Asp	Ile	Ile	Ser	Gly	Ala	His	Trp	G1y	Val	Met	Phe	
			345					350					355			
ggc	ctg	gcc	tac	ttc	tct	atg	cag	gga	gcg	tgg	gcg	aag	gtc	gtt	gtc	1459
Gly	Leu	Ala	Tyr	Phe	Ser	Met	G1n	Gly	Ala	Trp	Ala	Lys	: Val	Val	Val	
		360)				365					370)			
atc	cto	cte	g ctg	gcc	tct	ggg	gte	gac	gcg	tac	aco	acc	ace	g act	ggg	1507
Ile	Let	ı Let	ı Leı	ı Ala	Ser	Gly	Val	. Asp	Ala	Tyr	Thi	Thi	Thr	Thi	Gly	
	378	5				380)				388	5				
agc	gct	t gc	t ggg	g cgo	act	t acc	agt	t ago	c ctg	g gco	ag	c gc	c tto	c to	c cct	1555
															r Pro	
390					39					400					405	

ggc	gct	cgg	cag	aac	att	cag	ctc	att	aat	acc	aat	ggt	agc	tgg	cac	1603
Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr	Asn	Gly	Ser	Trp	His	
				410					415			٠		420		
atc	aac	cgc	acc	gcc	ctg	aat	tgc	aac	gat	tcc	ttg	cac	acc	ggc	ttc	1651
Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser	Leu	His	Thr	G1y	Phe	
		•	425		•		•	430					435		٠.	· . · · · .
				٠,٠				:				٠	· ·		٠	
ttc	acg	gcc	ctg	ttc	tac	atc	cat	aag	ttc	aac	tcg	tcg	gga	tgt	ccc	1699
Phe	Thr	Ala	Leu	Pĥe	Tyr	Ile	His	Lys	Phe	Asn	Ser	Ser	Gly	Cys	Pro	
		440	. •		·.		445					450				
gag	cgc	ctg	tcc	gcc	tgt	cgc	aac	atc	gag	gac	ttc	cgg	ata	gga	tgg	1747
Glu	Arg	Leu	Ser	Ala	Cys	Arg	Asn	Ile	G1u	Asp	Phe	Arg	Ile	G1y	Trp	
	455					460					465					
			•													
															agg	1795
Gly	Ala	Let	ı Gln	Tyr	Asp	Asp	Asn	Val	Thr	Asn	Pro	Glu	ı Asp	Met	Arg	
470					475	5				480					485	
cca	tat	tge	c tgg	g cac	tac	c cca	сса	aaa	cag	tgt	ggc	gta	a gtc	ccc	gca	1843
Pro	Туз	Cy:	s Trp	His	з Туз	rPro	Pro	Lys	G1n	Cys	Gly	/ Val	l Val	Pro	Ala	
				490)				495	5				500)	
ggg	aco	c gt	g tgo	c ggo	c cca	a gtg	g tac	tgt:	tto	acc	c cct	t ag	c ccg	g gtg	gta	1891
Gly	Th	r Va	1 Cys	s Gl	y Pr	o Val	l Tyı	c Cys	Phe	e Thi	Pro	o Se	r Pro	Va]	Val	
			50	5				510)				515	5		
~+~		0 00	a 20	ი და.	t. ao	a ct	t ggs	a gt.e	i cc.	t act	t ta	c ac	g tgg	g gga	a gag	1939

57/117

Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr Tyr Thr Trp Gly Glu 520 525 530

aat gag aca gat gtc ttc cta ttg aac agc acc cga cca ccg tcg ggg 1987
Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Ser Gly
535 540 545

tca tgg ttt ggc tgc acg tgg atg aac tcc act ggc ttc acc aag acc 2035 Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr 550 555 560 565

tgc ggc gca cca ccc tgc cgc act aga gct gac ttc aat acc agc aca 2083

Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Thr Ser Thr

570 575 580

gat ctg ttg tgc ccc acg gac tgt ttt aga aaa cat cct gaa gcc act 2131
Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr
585 590 595

tac atc aaa tgt ggt tcc ggg cct tgg ctc acg cca aag tgt ctg gtt 2179

Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val

600 605 610

gac tac ccc tac agg ctc tgg cat tac cct tgc aca gtc aat tac tcc 2227

Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Tyr Ser

615 620 625

acc ttc aag atc agg atg tat gtg ggg gga gtt gag cac agg ctc atg 2275
Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Met

2611

755

630					635					640					645	
gcc	gcg	tgc	aat	ttc	act	cgt	ggg	gat	cgc	tgc	aac	ttg	gag	gat	agg	2323
								Asp								
				650			·		655					660		
gac	aga	agt	caa	cag	act	cct	cţg	ttg	cac	tcċ	acc	acg	gaa	tgg	gcc	2371
Asp	Arg	Ser	Gln	Gln	Thr	Pro	Leu	Leu	His	Ser	Thr	Thr	Glu	Trp	Ala	·
•	• • •	·· ;	665		<u>.</u>		; ·	670	. • •	. ••		ξ.	675):		
_ 1 1	++~	000	7				•	ttg	•							2419
								Leu		•						
TIE	Leu	680	Cys	per	1110	001	685	200		-		690				
		000														
ctc	cac	ctc	cac	caa	aat	atc	gtg	gac	gta	caa	tat	atg	tat	ggc	ctg	2467
								Asp								
	695					700					705					
tca	cct	gcc	cto	aca	ı caa	tat	atc	gtt	cga	tgg	gag	tgg	g gta	gta	ctc	2515
Ser	Pro	Ala	Let	ı Thr	Gln	Tyr	Ile	· Val	Arg	Trp	Glu	Tr	Val	. Val	Leu	
710)				715	;				720)				725	
tta	tto	cte	g cto	c cta	a gcg	gac	gcc	agg	gto	tgo	gco	tg	c ttg	g tgg	g atg	2563
Leu	ı Phe	e Leu	ı Leı	ı Lei	u Ala	a Asp	Ala	a Arg	y Val	l Cys	s Ala	а Су	s Leu	ı Trp	Met	
				730	0				738	5				740)	

ctc atc ttg ctg ggc caa gcc gaa gca gca ctg gag aag ctg gtc gtc

Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val

745

750

ttg	cac	gct	gcg	agc	gca	gct	agc	tgc	aat	ggc	ttc	ctg	tat	ttt	gtc	2659	
Leu	His	Ala	Ala	Ser	Ala	Ala	Ser	Cys	Asn	Gly	Phe	Leu	Tyr	Phe	Val		
		760					765					770				•	
atc	ttt	ctc	gtg	gct	gct	tgg	cac	atç	aag	ggt	agg	gtg	gtc	ccc	ttg _.	2707	
Ile	Phe	Leu	Val	Ala	Ala	Trp	His	Ile	Lys	Gly	Arg	Val	Val	Pro	Leu		
	775					780					7,85		•		÷		
	•	•	<i>.</i>			•	٠	<i>:</i>	6	,	· . ·	··· ;					•
gct	gct	tat	tcc	ctt	act	ggc	ctg	tgg	ccg	ttc	tgc	cta	ctg	ctc	cta	2755	
Ala	Ala	Tyr	Ser	Leu	Thr	Gly	Leu	Trp	Pro	Phe	Cys	Leu	Leu	Leu	Leu		
790					795					800					805		
							•						•				
gca	ctg	ccc	cag	cag	gct	tac	gcc	tat	gat	gca	tct	gtg	cac	gga	cag	2803	
Ala	Leu	Pro	G1n	G1n	Ala	Tyr	Ala	Tyr	Asp	Ala	Ser	Val	His	G1y	Gln		
				810					815					820)		
gtg	ggc	gcg	gct	ttg	cta	gta	ctg	att	acc	ctc	ttt	aca	ctc	acc	ccg	2851	
Val	Gly	Ala	Ala	Leu	Leu	Val	Leu	Ile	Thr	Leu	Phe	Thr	Leu	Thr	Pro		
			825	,				830					835	5			
ggg	tat	aag	acc	ctt	cto	ago	cag	tcc	ctg	tgg	tgg	g ttg	tgo	tat	tctc	2899	
Gly	Tyr	Lys	Thr	Leu	Leu	ı Sei	Glr	Ser	Leu	Trp	Tr	Leu	с Суз	з Туз	c Leu		
		840)				845	5				850)				
ctg	aco	cta	g gcg	g gaa	a acc	at	g gto	cag	g gag	tgg	gca	a cca	a tc	c at	g cag	2947	
Leu	Thi	. Le	ı Ala	a Glu	ı Thi	. Me	t Vai	l G1r	ı Glu	ı Trp	Ala	a Pro	Se	r Me	t Gln		
	858	5				86	0				86	5					

							_ 4 _	-+-	+~~	~~~	acc	200	ata	+++	toc	2995
			ggc													2000
Ala .	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ile	Trp	Ala	Ala	Thr	lle	Phe		
870					875				• . (880					885	
ccg	ggc	gta	gtg	ttt	gac	ata	acc	aag	tgg	ctc	tta	gcg	gtg	ctt	ggg	3043
Pro	Gly	Val	Val	Phe:	Asp	Ile	Thr	Lys	Trp	Leu	Leu	Ala	Val	Leu	Gly	
٠.		· ·		890		- 30			895		•			900		
			-			•	: 4									
a a t	aat	tac	ctc	cta	้	øøt.	gct	ttg	acg	cgc	gtg	cċa	tat	ttc	gtc	3091
			Leu												200	
Pro	GTÀ	lyr			VI B	GIY	AIG		1 111	111 6	, 41		915			
			905					910					510			
											,					2120
															gcg	3139
Arg	Ala	His	Ala	Leu	Leu	Arg	Met	Cys	Thr	Met	Val	Arg	His	Let	Ala	
		920)				925					930)			
ggg	ggt	agg	g tac	gto	cag	atg	gcg	cta	tta	gcc	ctt	ggo	agg	g tgg	g act	3187
Gly	Gly	Arg	g Tyr	· Val	G1n	Met	Ala	Leu	Leu	Ala	ı Leı	ı Gly	/ Arg	g Trị	o Thr	
	935					940					945					
aac	act	t tar	r ato	tat	. gac	e cac	cto	acc	cct	ats	g to	g ga	t tgg	g gc	t gct	3235
															a Ala	
		r 1 y 1	C 116	÷ 1 y 1			Lec	. 1111	. 110	960		,	P1		965	
950					959	כ				900	J .				300	,
																0000
															c agt	
Ser	G1:	y Le	u Ar	g As	p Le	u Ala	a Val	l Ala	a Val	l G1	u Pr	o Il	e Il	e Ph	e Ser	
				97	0				97	5				98	0	
CCS	at at	g ga	g aa	g aa	a gt	c at	c gt	t tg	g gg	a gc	g ga	g ac	g gc	t go	g tg	3331

61/117

Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala Glu Thr Ala Ala Cys
985 990 995

ggg gac atc ttg cac gga ctt ccc gtg tcc gcc cga ctc ggt cgg gag 3379

Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala Arg Leu Gly Arg Glu

1000 1005 1010

atc ctc ctt ggc cca gct gat ggc tac acc tcc aag ggg tgg aag ctt 3427

Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser Lys Gly Trp Lys Leu

1015 1020 1025

ctc gcc ccc atc acc gct tac gcc cag cag aca cga ggt ctc ttg ggc 3475 Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly 1030 1035 1040 1045

tct ata gtg gtg agc atg acg ggg cgt gac aag aca gaa cag gcc ggg 3523 Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys Thr Glu Gln Ala Gly 1050 1055 1060

gag gtc caa gtc ctg tcc aca gtc act cag tcc ttc ctc gga aca tcc 3571
Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser Phe Leu Gly Thr Ser

1065 1070 1075

att tcg ggg gtc tta tgg act gtt tac cac gga gct ggc aac aag aca 3619

Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr

1080 1085 1090

cta gcc ggc tcg cgg ggc ccg gtc acg cag atg tac tcg agc gcc gag 3667 Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu

ggg gac ttg gtc ggg tgg ccc agc cct cct ggg acc aaa tct ttg gag Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu ccg tgt acg tgt gga gcg gtc gac ctg tat ttg gtc acg cgg aac gct Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala gat gtc atc ccg gct cga aga cgc ggg gac aag cgg gga gcg ctg ctc Asp Val Ile Pro Ala Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu tcc ccg aga ccc ctt tcg acc ttg aag ggg tcc tcg ggg gga cct gtg Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val ctt tgc cct agg ggc cac gct gtc gga atc ttc cgg gca gct gtg tgc Leu Cys Pro Arg Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys tct cgg ggt gtg gct aag tcc ata gat ttc atc ccc gtt gag acg ctc Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu

gac atc gtc acg cgg tct ccc acc ttt agt gac aac agc aca cca cca 4003

Asp Ile Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro

1210 1215 1220

gct .	gtg	ccc	cag	acc	tat	cag	gtg	ggg	tac	ttg	cac	gcc	ccc	act	ggc	4051
Ala	Val	Pro	Gln	Thr	Tyr	Gln	Val	Gly	Tyr	Leu	His	Ala	Pro	Thr	Gly	
		2	1225				1	230				3	235			
agt	gga	aaa	agc	acc	aag	gtc	ccc	gtc	gcg	tac	gcc	gcc	cag	ggg	tat	4099
Ser	Gly	Lys	Ser	Thr	Lys	Val	Pro	Val	Ala	Tyr	Ala	Ala	Gln	G1y	Tyr	
		1240			ě	j	1245	•	,		.]	1250	•		-	•
			: • •			· :			. •	•		. ,	-		. ,	*
aaa	gtg	ctg	gtg	ctc	aat	ccc	tcg	gtg	gct	gcc	acc	ctg	gga	ttt	ggg	4147
Lys	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	
. 1	1255					1260				;	1265					
						•										
gcg	tac	ttg	tcc	aag	gca	cat	ggc	atc	aac	ccc	aac	att	agg	act	gga	4195
															Gly	
1270					1275					1280					1285	
gtc	aga	act	t gtg	g acg	acc	ggg	gag	ccc	att	aca	tac	tco	ace	tai	t ggt	4243
															r Gly	
				1290)				1295					130	O	
aaa	tto	cto	c gc	c gat	ggg	g ggo	tgo	gca	ggo	ggc	gcc	ta	t gad	c at	c atc	4291
															e Ile	
			130					1310					131			
ata	tge	c ga	t ga	a tgo	c ca	c tc	t gtį	g gat	t gc1	t acc	c act	t at	t ct	c gg	c atc	4339
															y Ile	
		132					132					133				

ggg	aca	gtc	ctt	gac	caa	gca	gag	aca	gcc	ggg	gtc	agg	cta	act	gta	4387
Gly	Thr	Val	Leu	Asp	G1n	Ala	Glu	Thr	Ala	Gly	Val	Arg	Leu	Thr	Val	
	1335				1	340				1	345					
ctg	gcc	acg	gcc	acg	ccc	ccc	ggg	tcg	gtg	aca	acc	ccc	cat	ccc.	aat	4435
Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr	Thr	Pro	His	Pro	Asn	
135	0 .		• • •		1355		••••	· , .	1	1360				. 1	1365	
:		•	•	•				•	• •	٠.			•			
ata	gag	ġag	gta	gcc	ctc	ġga	cag	gag	ggt	gag	atc	ccc	ttc	tat	ggg	4483
Ile	Glu	Glu	Val	Ala	Leu	Gly	Gln	Glu	Gly	Glu	Ïle	Pro	Phe	Tyr	Gly	
				1370]	1375					1380		
							•									
agg	gcg	ttt	ccc	ctg	tct	tac	atc	aag	gga	ggg	agg	cac	ttg	att	ttc	4531
Arg	Ala	Phe	Pro	Leu	Ser	Tyr	Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	
			1385	;				1390					1395	;		
tgo	cac	tca	aag	g aaa	aag	tgt	gac	gag	ctc	gca	acg	gco	ctt	cgg	ggc	4579
Cys	His	Ser	Lys	s Lys	Lys	Cys	Asp	Glu	Leu	Ala	Thr	Ala	ı Let	ı Arg	Gly	
		1400)				1405	5				1410)			
at	g ggo	ttg	g aad	c gct	t gtg	g gca	tat	tac	aga	ggg	; ttg	gao	c gto	c tcc	ata	4627
Me	t Gly	y Let	ı Ası	n Ala	a Val	Ala	а Туі	Tyr	Arg	Gly	Leu	ı Ası	y Val	l Ser	· Ile	
	141	5			,	1420)				1425	5				
at	a cc	a ac	t ca	a gga	a gat	t gts	g gt	g gto	gtt	gco	aco	ga	c gc	c cto	atg	4675
11	e Pr	o Th	r Gl	n Gl	y Ası	y Va	l Va	l Val	. Val	l Ala	a Thi	r As	p Al	a Leu	ı Met	
14	30				143	5				1440)				1445	
ac	g gg	g ta	t ac	t gg	a ga	c tt	t ga	c tco	gt	gate	c ga	c tg	c aa	c gta	a gcg	4723

65/117

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala 1450 1455 1460

- gtc acc cag gcc gta gac ttc agc ctg gac ccc acc ttc act ata acc 4771

 Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr

 1465 1470 1475
- aca cag act gtc ccg caa gac gct gtc tca cgt agt cag cgc cga ggg 4819

 Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly

 1480 1485 1490
- cgc acg ggt aga gga aga ctg ggc att tat agg tat gtt tcc act ggt 4867
 Arg Thr Gly Arg Gly Arg Leu Gly Ile Tyr Arg Tyr Val Ser Thr Gly
 1495 1500 1505
- gag cga gcc tca gga atg ttt gac agt gta gta ctc tgt gag tgc tac 4915 Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr 1510 1515 1520 1525
- gac gca gga gct gct tgg tat gag ctc tca cca gtg gag acg acc gtc 4963
 Asp Ala Gly Ala Ala Trp Tyr Glu Leu Ser Pro Val Glu Thr Thr Val
 1530 1535 1540
- agg ctc agg gcg tat ttc aac acg cct ggc ttg cct gtg tgc cag gac 5011

 Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp

 1545 1550 1555
- cac ctt gag ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac 5059 His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp

PCT/JP2003/015038

gct cat ttc ctt tcc cag aca aag cag tcg ggg gaa aat ttc gca tac Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Ala Tyr tta gta gcc tat cag gcc aca gtg tgc gcc agg gcc aaa gcg ccc ccc Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ála Lys Ala Pro Pro ccg tcc tgg gac gtc atg tgg aag tgc ttg act cga ctc aag ccc acg Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr Arg Leu Lys Pro Thr ctt gtg ggc cct aca cct ctc ctg tac cgt ttg ggc tct gtt acc aac Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ser Val Thr Asn gag gtc acc ctt aca cac ccc gtg aca aaa tac atc gcc aca tgc atg Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr Ile Ala Thr Cys Met caa gct gac ctc gag gtc atg acc agc acg tgg gtc ctg gct ggg gga Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp Val Leu Ala Gly Gly gtc tta gca gcc gtc gcc gcg tat tgc tta gcg acc ggg tgt gtt tcc Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala Thr Gly Cys Val Ser

atc att ggc cgt tta cac atc aac cag cga gct gtc gtc gct ccg gac Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala Val Val Ala Pro Asp aag gag gtc ctc tat gag gct ttt gat gag atg gag gaa tgt gcc tcc Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Cys Ala Ser · 1705 aga gcg gct ctc ctt gaa gag ggg cag cgg ata gcc gag atg ctg aag Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile Ala Glu Met Leu Lys tcc aag atc caa ggc tta ttg cag caa gcc tct aaa cag gcc cag gac Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser Lys Gln Ala Gln Asp ata caa ccc gct gtg caa gct tcg tgg ccc aag atg gag caa ttc tgg Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys Met Glu Gln Phe Trp gcc aaa cat atg tgg aac ttc ata agc ggc att cag tac ctc gca gga Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly ctg tca aca ctg cca ggg aac cct gct gtg gct tcc atg atg gca ttc Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala Ser Met Met Ala Phe

agc	gcc	gcc	ctc	acc	agt	ccg	ttg	tca	act	agc	acc	acc	atc	ctt	ctt	5779
Ser	Ala	Ala	Leu	Thr	Ser	Pro	Leu	Ser	Thr	Ser	Thr	Thr	Ile	Leu	Leu	
•	1	1800				1	805				1	1810				
aac	att	ctg	ggg	ggc	tgg	ctg	gcg	tcc	caa	att	gcg	cca	ccc	gcg	ggg	5827
Asn	Ile	Leu	Gly	Gly	Trp	Leu				-		•	Pro	Ala	Gly	
1815 1820 1825															• • •	
			···					٠.	•		٠.		•			
gcc	act	ggc	ttt	gtt	gtc	agt	ggc	ctg	gtg	gga	gct	gct	gtt	ggc	agc	5875
Ala	Thr	Gly	Phe	Val	Val	Ser	Gly	Leu	Val	Gly	Ala	Ala	Val	Gly	Ser	
1830)			•	1835		•			1840			-	·	1845	
ata	ggc	ttg	ggt	aaa	gtg	ctg	gtg	gac	atc	ctg	gca	ggg	tat	ggt	gcg	5923
Ile	G1y	Leu	Gly	Lys	Val	Leu	Val	Asp	Ile	Leu	Ala	Gly	Tyr	G1y	Ala	
				1850					1855					1860		•
ggc	att	tcg	ggg	gcc	ctc	gtc	gcg	ttt	aag	atc	atg	tct,	ggo	gag	aag	5971
G1y	Ile	Ser	Gly	Ala	Leu	ı Val	Ala	Phe	Lys	Ile	Met	Ser	G1y	r Glu	Lys	
			1865	;				1870					1875	5		
ccc	tcc	ate	g gag	gat	gto	ato	aac	ttg	ctg	cct	ggg	g att	ct	g tct	cca	6019
Pro	Ser	Met	t Glu	ı Asp	Val	l Ile	Asn	Leu	Leu	Pro	G13	y Ile	e Lei	ı Sei	r Pro	
1880					1885					1890						
ggt	gc	t ct	g gtg	g gtg	g gga	a gto	ato	tgo	gc	g gcc	at	t ct	g cg	c cg	c cat	6067
Gly	Ala	a Le	u Val	l Val	L Gl	y Val	l Ile	e Cys	s Ala	a Ala	a Il	e Let	Ar;	g Ar	g His	1
	189	5				1900)				190	5				
gte	gg	a cc	g gg	g gaa	a gg	c gc	g gto	c caa	a tg	g at	g aa	c ag	g ct	t at	c gcc	6118

69/117

	Ala	Ile	Leu	Arg	Asn	Met	Trp	Gln	Val	Ala	G1y	Glu	Gly	Pro	Gly	Val
	1925	J				1920	i				1915	:			0	191
6163	gag	acg	gtg	tac	cac	act	cct	gcc	gtc	cac	aac	gga	aga	tcc	gct	ttc
	Glu	Thr	Val	Tyr	His	Thr	Pro	Ala	Val	His	Asn	Gly	Arg	Ser	Ala	Phe
		1940					1935	:				1930				
		•		*	•		• .	•		•	• .•	•		٠		
6211	ata	act	ctc	tct	ggc	ctt	ctġ	caa	acc	gtc	cgt	cag	tcg	gcg	gat	tcg
	Ile	Thr	Leu	Ser	Gly	Leu	Leu	Gln	Thr	Val	Arg	G1n	Ser	Ala	Asp	Ser
	, .		1955				•	1950	• • •		•	<i>:</i>	1945	• •		
										•						
6259	ccc	tgc	gat	gag	act	atc	tgg	aac	cac	ctt	aga	agg	ctc	cta	agt	act
	Pro	Cys	Asp	Glu	Thr	Ile	Trp	Asn	His	Leu	Arg	Arg	Leu	Leu	Ser	Thr
				1970	1				1965					1960	:	
6307	tgt	gtc	tgg	gac	tgg	gtg	gat	cgc	ctc	tgg	tcg	ggc	gcc	tgc	cca	atc
	Cys	Val	Trp	Asp	Trp	Val	Asp	Arg	Leu	Trp	Ser	Gly	Ala	Cys	Pro	Ile
					1985	1				1980	•				1975	-
6355	cca	ttc	ctg	aag	tcc	acc	ctg	tgg	aac	aag	ttt	gac	aca	cta	atc	acc
	Pro	Phe	Leu	Lys	Ser	Thr	Leu	Trp	Asn	Lys	Phe	Asp	Thr	Leu	Ile	Thr
	2005	2				2000	2				1995)	199
6403	ggc	aag	tac	ggg	aag	caa	tgc	tct	atc	ttt	ccc	ctc	ggc	cct	atg	aag
	Gly	Lys	Tyr	Gly	Lys	Gln	Cys	Ser	Ile	Phe	Pro	Leu	Gly	Pro	Met	Lys
		2020	:				2015	2				2010	:			
6451	gcc	ggc	tgc	ccc	tgc	cga	aca	acc	atg	atc	ggc	act	ggc	gcc	tgg	gtg
	Ala	Gl v	Cvs	Pro	Cvs	Aro	Thr	Thr	Met	Tle	G1 v	Thr	G1 v	Ala	Trp	Val

70/117

2025 2030 2035

aac atc tct ggc aac gtc cgc ttg ggc tct atg aga atc aca gga ccc 6499
Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro
2040 2045 2050

aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgt tat 6547

Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr

2055 2060 2065

aca gaa ggc cag tgc ttg ccg aaa ccc gcg tta aac ttc aag acc gcc 6595
Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu Asn Phe Lys Thr Ala
2070 2075 2080 2085

atc tgg aga gtg gcg gcc tca gag tac gcg gaa gtg acg cag cac gga 6643

Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly

2090 2095 2100

tca tat gcc tat ata aca ggg ctg acc act gac aac tta aaa gtc cct 6691 Ser Tyr Ala Tyr Ile Thr Gly Leu Thr Thr Asp Asn Leu Lys Val Pro 2105 2110 2115

tgc caa ctc ccc tct cca gag ttt ttc tct tgg gtg gac gga gta caa 6739 Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln 2120 2125 2130

atc cat agg tcc gcc ccc aca cca aag ccg ttt ttc cgg gat gag gtc 6787

Ile His Arg Ser Ala Pro Thr Pro Lys Pro Phe Phe Arg Asp Glu Val

2135 2140 2145

tcg	ttc	agc	gtt	ggg	ctc	aat	tca	ttt	gtc	gtc	ggg	tet	cag	CLL	CCC	0000
Ser	Phe	Ser	Val	G1y	Leu	Asn	Ser 1	Phe	Val	Val	Gly	Ser	G1n	Leu	Pro	
2150)			2	155				. 2	160				4	2165	
							·								•	
tgt	gac	cct	gag	ccc	gac	act	gag	gta	gţg	atg	tcc	atg	cta	aca	gac	6883
Cys	Așp	Pro	Glu	Pro	Asp	Thr	Glu	Val	Val	Met	Ser	Met	Leu	Thr	Asp	
			. :	2170			C.	. 2	2175	9		•	;	218 <u>0</u>		
	•			٠		٠.	•	• *						•		•
cca	tcc	cat	atc	acg	gcg	gag	gct	gca	gcg	cgg	cgt	tta	gcg	cgg	ggg	6931
Pro	Ser	His	Ile	Thr	Ala	Glu	Ala	Ala	Ala	Arg	Arg	Leu	Ala	Arg	Gly	
			2185				2	2190					2195			
tca	ccc	cca	tct	gag	gca	agc	tcc	tca	gcg	agc	cag	ctg	tcg	gcg	cca	6979
Ser	Pro	Pro	Ser	Glu	Ala	Ser	Ser	Ser	Ala	Ser	Gln	Leu	Ser	Ala	Pro	
		2200)				2205					2210)			
tcg	g ctg	g cga	a gcc	acc	tgc	acc	acc	cac	ggt	agg	acc	tat	gat	gt	g gac	7027
Ser	Let	ı Ar	g Ala	a Thr	Cys	Thr	Thr	His	Gly	Arg	Thr	Туг	. Asp	Va:	l Asp	
	2215	5				2220)				2225	;				
at	g gtį	g ga	t gc	c aac	c cta	g tto	atg	ggg	ggg	ggc	gte	g ati	t cg	gat	a gag	7075
															e Glu	
223					223					2240					2245	
tc	t ga	g tc	c aa	a gt	g gt	c gt	t ctg	g ga	c tco	cte	c ga	c tc	a at	g ac	c gag	7123
															r Glu	
				225					225					226	•	

gaa 1	gag	ggc	gac	ctt	gag	cct	tca	gta	cca	tcg	gag	tat	atg	ctc	ccc	7171
Glu (Glu	Gly	Asp	Leu	G1u	Pro	Ser	Val	Pro	Ser	Glu	Tyr	Met	Leu	Pro	
		2	2265				2	270				2	275			
			-											•		
agg	aag	agg	ttc	cca	ccg	gcc	tta	ccg	gct	tgg	gcg	cgg	cct	gat	tac	7219
Arg	Lys	Ārg	Phe	Pro	Pro	Äla	Leu	Pro	Ala	Trp	Ala	Arg	Pro	Asp	Tyr	
		2280		٠	•	. 2	285			٠.	2	2290				
۶.			•						٠.	-						
aac	cca	ccg	ctt	gtg	gaa	tcg	tgg	aag	agg	cca	gat	tac	caa	сса	ccc	7267
						Ser										•
	2295					2300					2305		•			
act	gtt	gcg	ggc	tgt	gct	ctc	ccc	ccc	ccc	aaa	aag	acc	ccg	ace	g cct	t 7315
						Leu										
2310					2315					2320					232	
cct	cca	agg	aga	. cgc	cgg	aca	gtg	ggt	ctg	agc	gag	ago	acc	ata	a gga	a 7363
						Thr										
110	110	, ,,,,	,	2330					2335					234		
				дос												
~a+	~ 00	. ata			r cte	g gcc	ato	aa e	tcc	ttt:	ggo	cas	z cc	c cc	c cc	a 7411
						ı Ala										
Asp	MIS	ı Le			I Dec	1 1110		2350			,,		235			
			234)				2000	,							
											. ~~		. «	o +o	o gra	c 7459
						t tco										
Ser	· G1:	y As	p Se	r Gl	y Le	u Sei			y Ala	a Asp	o Ala			р бе	er Gi	y
		236	0				236	5				237	U			
							٠									
gat	. cg	g ac	a cc	c cc	t ga	c ga	g tt	g gc	t ct	t tc	g ga	g ac	a gg	t to	t ac	c 7507

73/117

Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser Glu Thr Gly Ser Thr 2375 2380 2385

tcc tcc atg ccc ccc ctc gag ggg gag cct ggg gac cca gac ctg gag 7555
Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu
2390 2395 2400 2405

cct gag cag gta gag ctt caa cct cct ccc cag ggg ggg gag gca gct 7603 Pro Glu Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Glu Ala Ala 2410 2415 2420

ccc ggc tcg gac tcg ggg tcc tgg tct act tgc tcc gag gag gat gac 7651

Pro Gly Ser Asp Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp

2425 2430 2435

tcc gtc gtg tgc tgc tcc atg tca tat tcc tgg acc ggg gct cta ata 7699

Ser Val Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile

2440 2445 2450

act cct tgt agc ccc gaa gag gaa aag ttg cca att aac tcc ttg agc 7747

Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Ser Leu Ser

2455 2460 2465

aac tcg ctg ttg cga tac cat aac aag gta tac tgt act aca tca aag 7795
Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys
2470 2475 2480 2485

agt gcc tca cta agg gct aaa aag gta act ttt gat agg atg caa gtg 7843 Ser Ala Ser Leu Arg Ala Lys Lys Val Thr Phe Asp Arg Met Gln Val

PCT/JP2003/015038

2490

2495

2500

ctc g	ac	gcc	tat	tat	gat	tca	gtc	tta	aag	gac	atc	aag	cta	gcg	gcc	7891
Leu A	.sp	Ala	Tyr	Tyr	Asp	Ser	Val	Leu	Lys	Asp	Ile	Lys	Leu	Ala	Ala	
		2	505				2	510				2	2515			
tcc a	lag	gtc	agc	gca	agg.	ctc	ctc	acc	tta	gag	gag	gcg	tgc	caa	ttg	7939
Ser L	ys	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu	G1u	Ala	Cys	G1n	Leu	
	2	520			٠.	2	525			.=	:	2530		•		• ():
	••		9	··· ·				• •	. •				4	•	• •	
acc c	ca	ccc	cac	tct	gca	aga	tcc	aag	tat	ggg	ttt	ggg	gct	aag	gag	7987
Thr P																
	35					2540					2545					
gtc o	cgc	agc	ttg	tcc	ggg	agg	gcc	gtc	aac	cac	atc	aag	tcc	gtg	tgg	8035
Val A																
2550	J				2555					2560					2565	
aag g	gac	ctc	ttg	gaa	gac	tca	caa	aca	cca	att	cct	aca	. acc	atc	atg	8083
Lys A																
2,2.				2570	•				2575					2580		
gcc a	ลลล	aat	ฮลฮ	øt.ø	ttc	tgc	gtg	gac	ccc	gcc	aag	gge	ggt	aaa	aaa	8131
Ala																
nia i	Lys		2585		1 110			2590					2595		•	
			2000					2000								
c	~~±	. ~ -			+	. +	00+		+.		at a	. 201	otr	tar	, pac	8179
															gag	51.0
Pro				ı ITE	e val	LIYI			Let	ı GIŞ	va.			L Oys	oru	
		2600)		•		2605)				2610	J			

													.	_ +	0007
															8227
Met	Ala	Leu	Tyr	Asp	Val	Thr	Gln	Lys	Leu	Pro	Gln	Ala	Val	Met	
615				2	620				2	2625					
gct	tct	tat	ggc	ttc	cag	tac	tcc	ccc	gct	cag	cgg	gtg	gag	ttt	8275
Ala	Ser	Tyr	Gly.	Phe	Gln	Tyr	Ser	Pro	Ala	G1n	Arg	Val	Glu	Phe	• •
	٠									•					
٠.	. 6	٠,									•		·.		
ttg	aag	gca	tgg	gcg	gaa	aag	aga	gac	cct	atg	ggt	ttt	tcg	tat	8323
					•										
acc	cga	tgc	ttt	gac	tca	acc	gtc	act	gag	aga	gac	atc	agg	act	8371
Thr	Arg	Cys	Phe	Asp	Ser	Thr	Val	Thr	Glu	Arg	Asp	Ile	Arg	Thr	
gag	tcc	ata	tac	cag	gcc	tgc	tcc	tta	ccc	gag	gag	gcc	cga	act	8419
G1u	Ser	·Ile	Tyr	Gln	Ala	Cys	Ser	Leu	Pro	Glu	G1u	Ala	Are	Thr	
ata	ı ca	c tog	ctg	act	; gag	aga	cto	tat	gte	g gga	ı ggg	g ccc	ate	g ttc	8467
ago	c aa	g gg	c cag	g to	c tgo	ggg	g tao	c agg	g cg	t tgo	c cg	c gc	c ag	c ggg	8515
	J	•													
	Met 615 gct Ala ttg Leu acc Thr gag Glu ata 2698	Met Ala 615 gct tct Ala Ser ttg aag Leu Lys acc cga Thr Arg gag tcc Glu Ser 2680 ata cac Ile His 2695 c agc aa	Met Ala Leu 615 gct tct tat Ala Ser Tyr ttg aag gca Leu Lys Ala acc cga tgc Thr Arg Cys 2665 gag tcc ata Glu Ser Ile 2680 ata cac tcg Ile His Ser 2695 c agc aag ggc	Met Ala Leu Tyr 615 gct tct tat ggc Ala Ser Tyr Gly ttg aag gca tgg Leu Lys Ala Trp 2650 acc cga tgc ttt Thr Arg Cys Phe 2665 gag tcc ata tac Glu Ser Ile Tyr 2680 ata cac tcg ctg Ile His Ser Leu 2695 agc aag ggc cag Ser Lys Gly Gli	met Ala Leu Tyr Asp 615 2 gct tct tat ggc ttc Ala Ser Tyr Gly Phe 2635 ttg aag gca tgg gcg Leu Lys Ala Trp Ala 2650 acc cga tgc ttt gac Thr Arg Cys Phe Asp 2665 gag tcc ata tac cag Glu Ser Ile Tyr Gln 2680 ata cac tcg ctg act Ile His Ser Leu Thr 2695 agc aag ggc cag tcc Ser Lys Gly Gln Ser	Met Ala Leu Tyr Asp Val 2615 2620 gct tct tat ggc ttc cag Ala Ser Tyr Gly Phe Gln 2635 ttg aag gca tgg gcg gaa Leu Lys Ala Trp Ala Glu 2650 acc cga tgc ttt gac tca Thr Arg Cys Phe Asp Ser 2665 gag tcc ata tac cag gcc Glu Ser Ile Tyr Gln Ala 2680 ata cac tcg ctg act gag Ile His Ser Leu Thr Glu 2695 agc aag ggc cag tcc tgc Ser Lys Gly Gln Ser Cys	Met Ala Leu Tyr Asp Val Thr 2615 get tet tat gge tte cag tae Ala Ser Tyr Gly Phe Gln Tyr 2635 ttg aag gca tgg gcg gaa aag Leu Lys Ala Trp Ala Glu Lys 2650 acc cga tge ttt gae tea acc Thr Arg Cys Phe Asp Ser Thr 2665 gag tee ata tae cag gce tge Glu Ser Ile Tyr Gln Ala Cys 2680 2685 ata cae teg etg act gag aga Ile His Ser Leu Thr Glu Arg 2695 2700 age aag gge cag tee tge gge Ser Lys Gly Gln Ser Cys Gly	Met Ala Leu Tyr Asp Val Thr Gln 2615 2620 gct tct tat ggc ttc cag tac tcc Ala Ser Tyr Gly Phe Gln Tyr Ser 2635 ttg aag gca tgg gcg gaa aag aga Leu Lys Ala Trp Ala Glu Lys Arg 2650 acc cga tgc ttt gac tca acc gtc Thr Arg Cys Phe Asp Ser Thr Val 2665 2670 gag tcc ata tac cag gcc tgc tcc Glu Ser Ile Tyr Gln Ala Cys Ser 2680 2685 ata cac tcg ctg act gag aga ctc Ile His Ser Leu Thr Glu Arg Leu 2695 2700 c agc aag ggc cag tcc tgc ggg tac I Ser Lys Gly Gln Ser Cys Gly Tyr	Met Ala Leu Tyr Asp Val Thr Gln Lys 2615 2620 get tet tat gge tte eag tac tee eec Ala Ser Tyr Gly Phe Gln Tyr Ser Pro 2635 ttg aag gea tgg geg gaa aag aga gac Leu Lys Ala Trp Ala Glu Lys Arg Asp 2650 2655 acc ega tge ttt gae tea acc gte act Thr Arg Cys Phe Asp Ser Thr Val Thr 2665 2670 gag tee ata tac eag gee tge tee tee Glu Ser Ile Tyr Gln Ala Cys Ser Leu 2680 2685 ata cac teg etg act gag aga ete tat Ile His Ser Leu Thr Glu Arg Leu Tyr 2695 2700 c age aag gge cag tee tge ggg tac aga a Ser Lys Gly Gln Ser Cys Gly Tyr Arg	Met Ala Leu Tyr Asp Val Thr Gln Lys Leu 2615 2620 2 gct tct tat ggc ttc cag tac tcc ccc gct Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala 2635 2640 ttg aag gca tgg gcg gaa aag aga gac cct Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro 2650 2655 acc cga tgc ttt gac tca acc gtc act gag Thr Arg Cys Phe Asp Ser Thr Val Thr Glu 2665 2670 gag tcc ata tac cag gcc tgc tcc tta ccc Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro 2680 2685 ata cac tcg ctg act gag aga ctc tat gtg Ile His Ser Leu Thr Glu Arg Leu Tyr Val 2695 2700 c agc aag ggc cag tcc tgc ggg tac agg cg Ser Lys Gly Gln Ser Cys Gly Tyr Arg Ar	Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro 1615 2620 2625 get tet tat gge tte cag tac tee eec get cag Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln 10 2635 2640 ttg aag gea tgg geg gaa aag aga gae eet atg Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met 2650 2655 acc ega tge ttt gae tea acc gte act gag aga Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg 2665 2670 gag tee ata tac eag gee tge tee tta eec gag Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu 2680 2685 ata cac teg etg act gag aga etc tat gtg gga Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly 2695 2700 2706 age aag gge eag tee tge ggg tac agg egt tge a Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys	Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro Gln 1615 2620 2625 get tet tat gge tte eag tae tee eec get eag egg Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg 2635 2640 ttg aag gea tgg geg gaa aag aga gae eet atg ggt Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly 2650 2655 ace ega tge ttt gae tea ace gte act gag aga gae Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp 2665 2670 gag tee ata tae eag gee tge tee tta eec gag gag Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu 2680 2685 2690 ata cae teg etg act gag aga etc tat gtg gga ggg Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly 2695 2700 2705 age aag gge eag tee tge ggg tae agg egt tge egg I Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg	Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro Gln Ala 2615 2620 2625 gct tct tat ggc ttc cag tac tcc ccc gct cag cgg gtg Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val 2635 2640 ttg aag gca tgg gcg gaa aag aga gac cct atg ggt ttt Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly Phe 2650 2655 acc cga tgc ttt gac tca acc gtc act gag aga gac atc Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile 2665 2670 2685 2690 ata cac tcg ctg act gag aga ctc tat gtg gag gcc Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro 2695 agg aag agg cag tcc tgc ggg tac agg cgt tgc cgc gcc a Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala	Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro Gln Ala Val 615 2620 2625 gct tot tat ggc ttc cag tac tcc ccc gct cag cgg gtg gag Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu 2635 2640 ttg aag gca tgg gcg gaa aag aga gac cct atg ggt ttt tcg Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly Phe Ser 2650 2655 2660 acc cga tgc ttt gac tca acc gtc act gag aga gac atc agg Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg 2665 2670 2675 gag tcc ata tac cag gcc tgc tcc tta ccc gag gag gcc cga Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg 2680 2685 2690 ata cac tcg ctg act gag aga ctc tat gtg gga ggg ccc atg Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met 2695 2700 2705 agc aag gag cag tcc tgc ggg tac agg cgt tgc cgc gcc ag a Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Se	gct tct tat ggc ttc cag tac tcc ccc gct cag cgg gtg gag ttt Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Phe 2635 2640 2645 ttg aag gca tgg gcg gaa aag aga gac cct atg ggt ttt tcg tat Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly Phe Ser Tyr 2650 2655 2660 acc cga tgc ttt gac tca acc gtc act gag aga gac atc agg act Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg Thr 2665 2670 2675 gag tcc ata tac cag gcc tgc tcc tta ccc gag gag gcc cga act Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg Thr 2680 2685 2690 ata cac tcg ctg act gag aga ctc tat gtg gga ggg ccc atg ttc Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met Phe 2695 2700 2705 agc aag gag cag tcc tgc ggg tac agg cgt tgc cgc gcc agc ggg a Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly

										•						
gtg	ctt	acc	act	agt	atg	ggg	aac	acc	atc	aca	tgc	tat	gta	aaa	gcc	8563
Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr	Cys	Tyr	Val	Lys	Ala	
			2	2730		•		4	2735					2740		
cta	gcg	gct	tgc	aag	gct	gcg	gġg	ata	att	gcg	ccc	acg	atg	ctg	gta	8611
Leu	Ala	Ala	Cys	Lys	Ala	Ala	G1y	Ile	Ile	Ala	Pro	Thr	Met	Leu	Val	
			2745		•	. ;		2750					2755			•
		. :			,				ر مرخ							
tgc	ggc	gac	gac	ttg	gtc	gtc	atc	tca	gaa	agc	cag	ggg	act	gag	gag	8659
Cys	Gly	Asp	Asp	Leu	Val	Va1	Ile	Ser	G1u	Ser	G1n	Gly	Thr	Ğlu	G1u	•
•	2	2760				2	2765				2	2770				
gac	gag	cgg	aac	ctg	aga	gcc	ttc	acg	gag	gct	atg	acc	agg	tat	tct	8707
Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala	Met	Thr	Arg	Tyr	Ser	
2	2775				:	2780				2	2785					
									•							
gcc	cct	cct	ggt	gac	ccc	ccc	aga	ccg	gaa	tat	gac	ctg	gag	cta	ata	8755
Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr	Asp	Leu	Glu	Leu	Ile	
2790)			4	2795				2	2800				2	2805	
						-										
aca	tct	tgt	tcc	tca	aac	gtg	tct	gtg	gca	ctt	ggc	cca	cag	ggc	cgc	8803
Thr	Ser	Cys	Ser	Ser	Asn	Val	Ser	Val	Ala	Leu	G1y	Pro	Gln	Gly	Arg	
			2	2810				2	2815				2	2820		
cgc	aga	tac	tac	ctg	acc	aga	gac	ccc	acc	act	tca	att	gcc	cgg	gct	8851
Arg	Arg	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr	Ser	Ile	Ala	Arg	Ala	
		2	2825				2	2830				2	2835			
gcc	tgg	gaa	aca	gtt	aga	cac	tcc	cct	gtc	aat	tca	tgg	ctg	gga	aac	8899

77/117

Ala Trp Glu Thr Val Arg His Ser Pro Val Asn Ser Trp Leu Gly Asn
2840 2845 2850

- atc atc cag tac gct cca acc ata tgg gtt cgc atg gtc ctg atg aca 8947

 Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr

 2855 2860 2865
- cac ttc ttc tcc att ctc atg gcc cag gac acc cta gac cag aac ctt 8995

 His Phe Phe Ser IIe Leu Met Ala Gln Asp Thr Leu Asp Gln Asn Leu

 2870 2885 2880 2885
- aac ttt gaa atg tac gga tcg gtg tac tcc gtg agt cct ctg gac ctc 9043

 Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Ser Pro Leu Asp Leu

 2890 2895 2900
- cca gcc ata att gaa agg tta cac ggg ctt gac gcc ttc tct ctg cac 9091
 Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Leu His
 2905 2910 2915
- aca tac act ccc cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa 9139

 Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys

 2920 2925 2930
- ctt ggg gcg cca ccc ctc aga gcg tgg aag agt cgg gcg cgt gca gtt 9187 Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser Arg Ala Arg Ala Val 2935 2940 2945
- agg gcg tcc ctc atc tcc cgt ggg ggg agg gcg gcc gtt tgc ggt cgg 9235 Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala Ala Val Cys Gly Arg

2950 2955 2960 . 2965

tac ctc ttc aac tgg gcg gtg aag acc aag ctc aaa ctc act cct ttg 9283

Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu

2970 2975 2980

ccg gag gca cgc ctc ctg gat ttg tcc agt tgg ttt acc gtc ggc gcc 9331

Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala

2985 2990 2995

ggc ggg ggc gac att tat cac agc gtg tcg cgt gcc cga ccc cgc cta 9379

Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg Ala Arg Pro Arg Leu

3000 3005 3010

tta ctc ctt agc cta ctc cta ctt tct gta ggg gta ggc ctc ttc cta 9427

Leu Leu Leu Ser Leu Leu Leu Ser Val Gly Val Gly Leu Phe Leu

3015 3020 3025

ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct 9482
Leu Pro Ala Arg
3030

ggctagctgt gaaaggtccg tgagccgcat gactgcagag agtgccgtaa ctggtctctc 9662

9674

<210> 6

<211> 3033

<212> PRT

<213> Hepatitis C virus

<400> 6

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala

35
40
45

Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
50 55 60

Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp

85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro

100 105 110

Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
115 120 125

Gly Phe Ala Asp Leu Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu 130 135 140

Ser Gly Val Ala Ser Ala Leu Ala His Gly Val Arg Val Leu Glu Asp 145 150 155 160

Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile

				165					170					175	•
Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Ile	Thr	Thr	Pro	Val	Ser	Ala	Val
			180					185					190		
Gln	Val	Lys	Asn	Thr	Ser	Asn	Ala	Tyr	Met	Ala	Thr	Asn	Asp	Cys	Ser
		195					200					205			
Asn	Asp	Ser	Ilė	Thr	Trp	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Val	Pro
	210	:		-		215		٠.	•		220			٠	
Gly	Cys	Val	Pro	Cys	Glu	Lys	Met	Gly	Asn	Thr	Ser	Arg	Cys	Trp	Ile
225					230			•	٠.	235		·			240
Pro	Val	Ser	Pro	Asn	Val	Ala	Val	Arg	Gln	Pro	G1y	Ala	Leu	Thr	Arg
				245		•			250					255	
G1y	Leu	Arg	Thr	His	Ile	Asp	Met	Val	Val	Leu	Ser	Ala	Thr	Leu	Cys
			260					265					270		
Ser	Ala	Leu	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Gly	Val	Met	Leu	Ala	Ser
		275					280					285			
Gln	Met	Phe	Ile	Val	Ser	Pro	Gln	His	His	Trp	Phe	Val	Gln	Glu	Cys
	290					295		٠	•		300				
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	Ala	Ile	Thr	Gly	His	Arg	Met	Ala	Trp
305					310					315					320
Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Thr	Thr	Met	Ile	Leu	Ala	Tyr
				325					330					335	
Val	Met	Arg	Val	Pro	Glu	Val	Ile	Ile	Asp	Ile	Ile	Ser	Gly	Ala	His
			340					345					350		
Trp	Gly	Val	Met	Phe	Gly	Leu	Ala	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp
		355					360					365	٠		
Ala	Lys	Val	Val	Val	Ile	Leu	Leu	Leu	Ala	Ser	Gly	Val	Asp	Ala	Tyr
	370					375					380				
Thr	Thr	Thr	Thr	Gly	Ser	Ala	Ala	Gly	Arg	Thr	Thr	Ser	Ser	Leu	Ala
385					390					395					400

Ser	Ala	Phe	Ser	Pro	Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr
				405					410					415	
Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Alạ	Leu	Asn	Cys	Asn	Asp	Ser
			420					425					430		
Leu	His	Thr	Gly	Phe	Phe	Thr	Ala	Leu	Phe	Tyr	Ile	His	Lys	Phe	Asn
		435					440					445		•	
Ser	Ser	Gly	Cys	Pro	Glu	Arg	Leu	Ser	Ala	Cys	Arg	Asn	Tle	Glu	Asp
•	450			•		455	•	•			460				•
Phe	Arg	Ile	Gly	Trp	Gly	Ala	Leu	Gln	Tyr	Asp	Asp	Asn	Val	Thr	Asn
465	7		•	:	470		•	• •		475				•	480
Pro	Glu	Asp	Met	Arg	Pro	Tyr	Cys	Trp	His	Tyr	Pro	Pro	Lys	Gln	Cys
				485					490	•				495	
G1y	Val	Val	Pro	Ala	Gly	Thr	Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr
			500					505					510		
Pro	Ser	Pro	Val	Val	Val	G1y	Thr	Thr	Asp	Arg	Leu	G1y	Val	Pro	Thr
		515	i				520					525			
Tyr	Thr	Trp	Gly	Glu	Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr
	530	ļ				535					540)			
Arg	Pro	Pro	Ser	Gly	Ser	Trp	Phe	Gly	Cys	Thr	Trp	Met	. Asn	Ser	Thr
545					550					555					560
G1y	Phe	Thi	Lys	Thr	Cys	Gly	Ala	Pro	Pro	Cys	Arg	g Thr	Arg		Asp
				565			•		570					575	
Phe	e Asr	Th	c Sei	Thr	Asp	Leu	ı Leu	Cys	Pro	Thi	. Asr	Cys			g Lys
			580					585					590		
His	Pro	Gl:	u Ala	a Thi	Туз	c Ile	e Lys	cys Cys	Gly	r Sei	c Gly) Let	ı Thr
		59					600					608			
Pro	Lys	s Су	s Lei	u Val	L Ası	о Туі	rPro	туг	Are	g Lei	ı Tr	o His	з Тул	r Pro	o Cys
	610					615					62				
Thi	r Va	l As	n Ty:	r Se	r Th	r Phe	e Lys	s Ile	Ar	g Me	t Ty:	r Va	1 G1;	y Gl	y Val
								82	2/11	7					

625	•				630					635					640
Glu	His	Arg	Leu	Met	Ala	Ala	Cys .	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Cys
				645					650					655	
Asn	Leu	Glu	Asp	Arg	Asp	Arg	Ser	Gln	Gln	Thr	Pro	Leu	Leu	His	Ser
			660					665					670		
Thr	Thr	Glu	Trp	Ala	Ile	Leu	Pro	Cys	Ser	Phe	Ser	Asp	Leu	Pro	Ala
•	•	675	-	•		•	680			.		685		٠.	
Leu	Ser	Thr	Gly	Leu	Leu	His	Leu	His	G1n	Asn	Ile	Val	Asp	Val	G1n
•	690					695					700 _.				
Tyr	Met	Tyr	Gly	Leu	Ser	Pro	Ala	Leù	Thr	Gln	Tyr	Ile	Val	Arg	Trp
705					710					715					720
Glu	Trp	Val	Val	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Cys
				725					730					735	
Ala	Cys	Leu	Trp	Met	Leu	Ile	Leu	Leu	Gly	G1n	Ala	Glu	Ala	Ala	Leu
			740					745					750		
Glu	Lys	Leu	Val	Val	Leu	His	Ala	Ala	Ser	Ala	Ala	Ser	Cys	Asn	G1y
		755					760					765			
Phe	Leu	Tyr	Phe	Val	Ile	Phe	Leu	Val	Ala	Ala	Trp	His	Ile	Lys	Gly
	770	1				775					780				
Arg	Val	. Val	Pro	Leu	Ala	Ala	Tyr	Ser	Leu	Thr	Gly	Leu	Trp	Pro	
785					790					795					800
Cys	Leu	ı Lev	ı Leı	ı Leu	ı Ala	Leu	Pro	Gln	G1n	ı Ala	Tyr	Ala	Tyr	Asp	Ala
				805	5				810)				815	
Ser	Va]	His	s G1;	y Glr	ı Val	Gly	Ala	. Ala	Lei	ı Leu	ı Val	Let	ı Ile	Thr	Leu
			820	0				825	5				830)	
Phe	Thi	r Lei	ı Th	r Pro	o G13	y Tyr	Lys	Thr	Le	ı Leı	ı Sei	Glr	ı Ser	Leu	Trp
		838	5				840)				849	5		
Trp	Lei	и Су	s Ту	r Lei	u Lei	ı Thı	Lei	ı Ala	a Gl	u Thi	r Met	t Vai	l G1r	ı Glu	ı Trp
	85	0				858	5				860)			

Alal	Pro	Ser	Met	Gln	Ala	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ile	Trp	Ala
865					870					875					880
Ala '	Thr	Ile	Phe	Cys	Pro	Gly	Val	Val	Phe	Asp	Ile	Thr	Lys	Trp	Leu
				885					890					895	
Leu	Ala	Val	Leu	Gly	Pro	G1y	Tyr	Leu	Leu	Arg	Gly	Ala	Leu	Thr	Arg
			900					905	•				910	·	
Val	Pro	Tyr	Phe	Val	Arg	Alà	His	Ala	Leu	Leu	Arg	Met	Cys	Thr	Met
		915	•	·			920			. *		925			•
Val	Arg	His	Leu	Ala	G1y	Gly.	Arg	Tyr	Val	G1n	Met	Ala	Leu	Leú	Ala
٠	930	•		,		935	•	•	:		940	•	•		
Leu	Gly	Arg	Trp	Thr	Gly	Thr	Tyr	Ile	Tyr	Asp	His	Leu	Thr	Pro	Met
945				•	950					955					960
Ser	Asp	Trp	Ala	Ala	Ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu
				965					970					975	
Pro	Ile	Ile	Phe	Ser	Pro	Met	Glu	Lys	Lys	Val	Ile	Val	Trp	Gly	Ala
			980					985					990		
G1u	Thr	Ala	Ala	Cys	G1y	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala
		995	;				1000					1005			
Arg	Leu	Gly	Arg	g Glu	Ile	Leu	Leu	G1y	Pro	Ala	Asp	Gly	Tyr	Thr	Ser
:	1010)				1015	i				1020	•			
Lys	Gly	Trp	Lys	s Leu	Leu	ı Ala	Pro	Ile	Thr	Ala	Tyr	Ala	G1n	Gln	Thr
102	5				1030)				1035	,				1040
Arg	G13	, Lei	ı Lev	ı Gly	Ser	· Ile	val	Val	Ser	. Met	Thr	Gly	Arg	Asp	Lys
				1045	5				1050)				1055	;
Thr	Glı	ı Glı	n Ala	a Gly	7 G11	ı Val	l G1r	ı Val	Leu	ı Sei	Thi	· Val	Thr	G1r	Ser
			1060	0				1065	5				1070)	
Phe	Let	ı G1;	y Th	r Sei	r Ile	e Sei	r Gly	/ Val	Let	u Tr	Thi	· Val	Туз	His	Gly
		107	5				1080)				1085	5		
Ala	G1	y Ası	n Ly	s Thi	r Lei	u Ala	a Gly	y Sei	Arı	g Gl	y Pro	o Val	l Thi	c Glı	n Met
								84	4/11′	7	_				

10	090				1	095	٠			1	100				
Tyr :	Ser	Ser	Ala	G1u	G1y	Asp	Leu	Val	Gly	Trp	Pro	Ser	Pro	Pro (Gly
1105				1	.110				1	115				1	120
Thr	Lys	Ser	Leu	Glu	Pro	Cys	Thr	Cys	Gly	Ala	Val	Asp	Leu	Tyr	Leu
			1	125				1	130				1	135	
Val	Thr	Arg	Asn	Ala	Asp	Val	Ile	Pro	Ala	Arg	Arg	Arg	G1y	Asp	Lys
•	•	•:]	1140				· " 1	145	•		•	.1	150	-	
Arg	Gly	Ala	Leu	Leu	Ser	Pro	Arg	Pro	Leu	Ser	Thr	Leu	Lys	Gly	Ser
	• •	1155				-	1160					1165			
Ser	Gly	G1y	Pro	Val	Leu	Cys	Pro	Arg	Gly	His	Ala	Val	Gly	Ile	Phe
1	170]	175					1180		•		
Arg	Ala	Ala	Val	Cys	Ser	Arg	Gly	Val	Ala	Lys	Ser	Ile	Asp	Phe	Ile
1185	5				1190				:	1195	•			1	.200
Pro	Val	Glu	Thr	Leu	Asp	Ile	Val	Thr	Arg	Ser	Pro	Thr	Phe	Ser	Asp
				1205					1210				d	1215	
Asn	Ser	Thr	Pro	Pro	Ala	Val	Pro	G1n	Thr	Tyr	Gln	Val	Gly	Tyr	Leu
			1220			•		1225					1230		
His	Ala	Pro	Thr	Gly	Ser	Gly	Lys	Ser	Thr	Lys	Va1	Pro	Val	Ala	Tyr
		1235					1240					1245			
Ala	Ala	G1n	Gly	Tyr	Lys	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala
3	1250)				1255					1260)			
Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Leu	Ser	Lys	Ala	His	Gly	Ile	Asn	Pro
126	5				1270)				1275	5				1280
Asn	Ile	Arg	Thr	Gly	v Val	Arg	Thr	Val	Thr	Thr	Gly	Glu	Pro	Ile	Thr
				1285	5				1290)				1295	
Tyr	Ser	Thi	Tyr	G13	/ Lys	Phe	Leu	Ala	Asp	Gly	Gly	r Cys	Ala	Gly	Gly
			1300)				1305	5				1310	1	
Ala	Tyı	: Ası	o Ile	e Ile	e Ile	e Cys	s Asp	Glu	ı Cys	His	s Sei	· Val	Asp	Ala	Thr
		1319	5				1320)				1325	5		

Thr Ile	Leu G	ly :	Ile	Gly	Thr	Val	Leu .	Asp	Gln .	Ala	Glu	Thr .	Ala (Gly
1330					335					340				
Val Arg	Leu T	hr '	Val :			Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val '	Thr
1345				350					355					360
Thr Pro	His F	ro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Gly	Gln	Glu	Gly	Glu
			365					370					375	
Ile Pro	Phe 7	ſyr	Gly	Arg	Álá	Phe	Pro	Leù	Ser	Tyr	Ile	Lys	Gly	Gly
		380					1385					390		•
Arg His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu	Ala
1	L395	•			:	1400		:		` :	1405			<i>:</i> .
Thr Ala	Leu .	Arg	Gly	Met	Gly	Leu	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly
1410					1415]	L 4 20				
Leu Asp	Val	Ser	Ile	Ile	Pro	Thr	Gln	Gly	Asp	Val	Val	Val	Val	Ala
1425			:	1430					1435				1	1440
Thr Asp	Ala	Leu	Met	Thr	-Gly	Tyr	·Tḥr	Gly	Asp	Phe	Asp	Ser	Val	Ile
			1445				,	1450					1455	
Asp Cys	Asn	Val	Ala	Val	Thr	G1n	Ala	Val	Asp	Phe	Ser	Leu	Asp	Pro
	1	460					1465					1470		
Thr Phe	Thr	Ile	Thr	Thr	Gln	Thr	Val	Pro	G1n	Asp	Ala	Val	Ser	Arg
	1475					1480)				1485			
Ser Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Arg	Leu	Gly	Ile	Tyr	Arg
1490)				1495	5				1500)			
Tyr Val	Ser	Thr	Gly	Glu	ı Arg	g Ala	a Ser	Gly	Met	Phe	e Asp	Ser	Val	Val
1505				1510)				1515	;				1520
Leu Cys	Glu	Cys	Tyr	. Asp	Ala	a Gl	y Ala	ı Ala	a Trp	Tyı	Glu	ı Leu	. Ser	Pro
			1525	5				1530)				1535	
Val Glu	ı Thr	Thr	· Val	l Ar	g Lei	ı Ar	g Ala	а Туз	r Phe	e Ası	n Thi	r Pro	Gly	Leu
		1540)				1545	5				1550)	
Pro Val	l Cys	Glı	n Asp	o Hi	s Le	u Gl	u Phe	e Tr	p Glu	Al:	a Vai	l Phe	: Thr	Gly
							. 80	6/11′	7					

	1	555				1	560				1	565			
Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	Gln	Thr	Lys	Gln	Ser (Gly
1	570				1	575				1	580				•
G1u	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Tyr	G1n	Ala	Thr	Val	Cys	Ala	Arg
1585				1	590				1	595				1	600
Ala	Lys	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Val	Met	Trp	Lys	Cys	Leu	Thr
				1605		•••.	: '	1	610	= .		• ,	1	615	
Arg	Leu	Lys	Pro	Thr	Leu	Vaĺ	Gly	Pro	Thr	Pro	Leu	Leu	Tyr	Arg	Leu
			1620	:			•	1625	• •			1	630		0
Gly	Ser	Val	Ťhr	Asn	Glu	Val	Thr	Leu	Thr	His	Pro	Val	Thr	Lys	Tyr
		1635				•	1640				;	1645			
Ile	Ala	Thr	Cys	Met	G1n	Ala	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp
3	1650					1655					1660				
Val	Leu	Ala	Gly	Gly	Val	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Cys	Leu	Ala
166	5				1670				(1675				1	.680
Thr	Gly	Cys	Val	Ser	Ile	Ile	Gly	Arg	Leu	His	Ile	Asn	Gln	Arg	Ala
				1685					1690					1695	
Val	Val	Ala	Pro	Asp	Lys	Glu	Val	Leu	Tyr	G1u	Ala	Phe	Asp	Glu	Met
			1700					1705					1710		
Glu	Glu	ι Суз	: Ala	a Ser	Arg	Ala	ı Ala	Leu	Leu	G1u	Glu	Gly	G1n	Arg	Ile
		1715					1720					1725		•	
Ala	Glu	ı Met	: Le	ı Lys	Ser	Lys	s Ile	e Gln	Gly	Let	ı Lev	Gln	Gln	Ala	Ser
	1730)		•		1739	5				1740)			
Lys	Glr	n Ala	a Gl	n Asp	Ile	e Gli	n Pro	Ala	Val	G1r	n Ala	Ser	Trp	Pro	Lys
174	5				1750)		•		1759	5				1760
Met	G11	u Gl	n Ph	e Tr	o Ala	a Lys	s His	s Met	Trp	Ası	n Phe	e Ile	Ser	Gly	Ile
				1769	5				1770)				1775	
G1r	ı Ty:	r Le	u Al	a Gl	y Le	ı Se	r Th	r Lei	ı Pro	G1;	y Ası	n Pro	Ala	a Val	Ala
			178	0				1789	5				1790)	

Ser	Met	Met	Ala	Phe	Ser	Ala	Ala	Leu	Thr	Ser	Pro	Leu	Ser	Thr	Ser
	1	795				1	.800				1	.805			
Thr	Thr	Ile	Leu	Leu	Asn	Ile	Leu	Gly	Gly	Trp	Leu	Ala	Ser	Gln	Ile
J	1810				1	815				1	.820				
Ala	Pro	Pro	Ala	G1y	Ala	Thr	Gly	Phe	Val	Val	Ser	Gly	Leu	Val	Gly
182	5			1	1830]	1835				, 1	840
Ala	Ala	Va1	G1y	Ser	Ile	Gly	Leu	Gly	Lys	Val	Leu	Val	Asp	Ile	Leu
·				1845				. 1	850					1855	*
Ala	Gİy	Tyr	G1y	Ala	G1y	Ile	Ser	Gly	Ala	Leu	Val	Ala	Phe	Lys	Ϊle
			1860		-		•	1865	` <u></u>	•		:	1870		
Mèt	Ser	G1y	Glu	Lys	Pro	Ser	Met	Glu	Asp	Val	Ile	Asn	Leu	Leu	Pro
		1875					1880					1885			
G1y	Ile	Leu	Ser	Pro	Gly	Ala	Leu	Val	Val	Gly	Val	Ile	Cys	Ala	Ala
	1890					1895					1900				
Ile	Leu	Arg	Arg	His	Val	G1y	Pro	Gly	Glu	Gly	Ala	Val	G1n	Trp	Met
190	5				1910	1				1915					1920
Asn	Arg	g Leu	Ile	e Ala	Phe	Ala	Ser	Arg	G1y	Asn	His	Val	Ala	Pro	Thr
				1925	;			. ·	1930)				1935	
His	туг	· Val	Thi	Glu	Ser	Asp	Ala	Ser	Gln	Arg	; Val	Thr	· Glr	ı Leu	Leu
			1940)			-	1945					1950)	
G13	7 Sei	Leu	ı Thi	r Ile	Th:	: Ser	Leu	ı Leu	Are	g Arg	Let	ı His	s Ası	n Trp	Ile
		1955	5				1960)				1965	5		
Thi	r Glı	ı Asp	Cy:	s Pro	o Ile	e Pro	Cys	s Ala	G13	7 Sei	Tr	Let	ı Arı	g Asp	Val
	1970	0				1978	5				1980)			
Tr	p As	p Tr	y Va	1 Cys	s Th	r Ile	e Let	ı Thr	. Ası	o Phe	e Lys	s Ası	n Tr	p Leu	Thr
19	85				199	0				199	5				2000
Se	r Ly	s Le	u Ph	e Pro	o Ly	s Me	t Pro	o Gly	/ Le	u Pro	o Pho	e Ile	e Se	r Cys	Gln
				200	5				201	0				2015	5
Ly	s Gl	у Ту:	r Ly	s Gl	y Va	1 Tr	p Ala	a Gl	y Th	r Gl	y Il	e Me	t Th	r Thi	Arg
						,			8/11						

		2	020				2	025				2	030		
Cys	Pro	Cys	G1y	Ala	Asn	Ile	Ser	Gly A	Asn	Val	Arg	Leu	Gly	Ser 1	Met
	2	2035				2	040				2	045			
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Cys	Met .	Asn	Thr	Trp	G1n	Gly	Thr	Phe
2	050				2	055				2	060				
Pro	Ile	Asn	Cys	Tyr	Thr	Glu	Gly	Gln .	Cys	Leu	Pro	Lys	Pro	Ala	Leu
2065	•	•			2070					2075					080
Asn	Phe	Lys	Thr	Ála	Ile	Trp	Arg	Val	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
		•		2085		. •	·-		090	•				2095	
Val	Thr	Gln	His	Gly	Ser	Tyr	Ala	Tyr	Ile	Thr	Gly	Lèu	Thr	Thr	Asp
			2100			•		2105					2110		
Asn	Leu	Lys	Val	Pro	Cys	G1n	Leu	Pro	Ser	Pro			Phe	Ser	Trp
		2115					2120					2125		•	
Val	Asp	Gly	Val	G1n	Ile	His	Arg	Ser	Ala	Pro	Thr	Pro	Lys	Pro	Phe
	2130)			.;	2135				:	2140				
Phe	Arg	Asp	Glu	Val	Ser	Phe	Ser	Val	G1y	Leu	Asn	Ser	Phe	Val	Val
214	5				2150					2155				2	2160
Gly	Ser	Gln	Leu	Pro	Cys	Asp	Pro	Glu	Pro	Asp	Thr	Glu	Val	Val	Met
				2165				:	2170				:	2175	
Ser	Met	Leu	ı Thr	Asp	Pro	Ser	His	Ile	Thr	Ala	Glu	Ala	Ala	Ala	Arg
			2180)				2185					2190		
Arg	Lei	ı Ala	a Arg	g Gly	Ser	Pro	Pro	Ser	Glu	Ala	Ser	Ser	Ser	Ala	Ser
		2199	5				2200)				2205			
Glr	Le	u Ser	c Ala	a Pro	Ser	Leu	ı Arg	, Ala	Thr	Cys	Thr	Thr	His	Gly	Arg
	221	0				2215	5				2220)			
Thi	Ty:	r As	p Va	l Asp	o Met	: Val	l Asp) Ala	Asr	ı Leu	Phe	Met	Gly	Gly	Gly
222	25				2230)				2235	5				2240
Val	l I1	e Ar	g Il	e Glu	u Sei	c Glu	u Sei	Lys	Va]	l Val	[Va]	Lei	ı Asp	Ser	Leu
				224	5				2250)				2255	•

Asp	Ser	Met	Thr	Glu	Glu	Glu	G1y	Asp	Leu	Glu	Pro	Ser	Val	Pro	Ser
		2	260				2	265				2	270		
Glu	Tyr	Met	Leu	Pro	Arg	Lys	Arg	Phe	Pro	Pro	Ala	Leu	Pro	Ala	Trp
	2	2275	•			2	280				2	285			
Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Lys	Arg	Pro
2	2290				2	2295			er e	2	2300				
Asp	Tyr	G1n	Pro	Pro	Thr	Val	Ala	Gly	Cys	Ala	Leu	Pro	Pro	Pro	Lys
230	5		•	2	2310				2	2315	0			2	2320
Lys	Thr	Pro	Thr	Pro	Pro	Pro	Arg	Arg	Arg	Arg	Thr	Val	Gly	Leu	Ser
,	÷	٠,		2325	٠.	• .	•	2	2330				. 2	2335	
Glu	Ser	Thr	Ile	Gly	Asp	Ala	Leu	Gln	G1n	Leu	Ala	Ile	Lys	Ser	Phe
			2340				2	2345				2	2350		
Gly	Gln	Pro	Pro	Pro	Ser	Gly	Asp	Ser	Gly	Leu	Ser	Thr	Gly	Ala	Asp
		2355		•		3	2360				;	2365			
Ala	Ala	Asp	Ser	Gly	Asp	Arg	Thr	Pro	Pro	Asp	Glu	Leu	Ala	Leu	Ser
	2370)	•		•	2375				Ġ	2380	•			
Glu	Thr	Gly	Ser	Thr	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	Pro	Gly
238	5				2390		•	•		2395					2400
Asp	Pro	Asp	Leu	Glu	Pro	G1u	Gln	Val	G1u	Leu	G1n	Pro	Pro	Pro	Gln
				2405	;		•		2410					2415	
G1y	Gly	Glu	ı Ala	ı Ala	Pro	G1y	Ser	Asp	Ser	Gly	Ser	Trp	Ser	Thr	Cys
•			2420)				2425					2430		
Ser	Glu	ı Glu	ı Asp	Asp	Ser	Val	Val	Cys	Cys	Ser	Met	Ser	Tyr	Ser	Trp
		2435	5				2440)				2445			
Thi	G1;	y Ala	a Lei	ı Ile	Thr	Pro	Cys	Ser	Pro	Glu	Glu	Glu	Lys	Leu	Pro
	2450)				2455	5				2460)			
Ile	e Ası	n Sei	r Lei	ı Sei	. Asr	ı Ser	Leu	ı Leu	ı Ar	g Tyr	His	. Asn	Lys	Val	Tyr
246					2470					2475					2480
		r Thi	r Se	r Lys	s Sei	c Ala	a Sei	Leu	ı Ar	g Ala	a Lys	s Lys	. Val	Thr	Phe
									0/117						

			2	2485				2	490				2	495	
Asp	Arg	Met	Gln	Val	Leu	Asp	Ala	Tyr	Tyr	Asp	Ser	Val	Leu	Lys	Asp
		2	2500				2	505				2	2510		
Ile	Lys	Leu	Ala	Ala	Ser	Lys	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	G]lu
	:	2515				2	2520				2	2525			
Glu	Ala	Cys	Gln	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	Ser	Lys	Tyr	Gly
. 2	2530	:			2	2535	•			. 2	2540			1	•
Phe	Gly	Ala	Lys	Glu	Val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His
2545	5 ,	 -		.2	2550	•	•		. 2	2555		· · · <u>-</u>		. 2	2560
Ile	Lys	Ser	Val	Trp	Lys	Ásp	Leu	Ĺeu	Glu	Asp	Ser	G1n	Thr	Pro	Ile
				2565					2570			**		2575	
Pro	Thr	Thr	Ile	Met	Ala	Lys	Asn	Glu	Val	Phe	Cys	Val	Asp	Pro	Ala
			2580				:	2585					2590		
Lys	G1y	G1y	Lys	Lys	Pro	Ala	Arg	Leu	Ile	Val	Tyr	Pro	Asp	Leu	Gly
		2595					2600				;	2605			
Val	Arg	; Val	Cys	G1u	Lys	Met	Ala	Leu	Tyr	Asp	Val	Thr	Gln	Lys	Leu
	2610)				2615					2620				
Pro	Gln	Ala	l Val	Met	Gly	Ala	Ser	Tyr	Gly	Phe	G1n	Tyr	Ser	Pro	Ala
262	5				2630					2635					2640
Gln	Are	y Val	. Gl u	ı Phe	Leu	Leu	Lys	Ala	Trp	Ala	G1u	Lys	Arg	Asp	Pro
			•	2645					2650)				2655	
Met	G13	7 Phe	e Ser	Tyr	Asp	Thr	Arg	Cys	Phe	Asp	Ser	Thr	Val	Thr	Glu
			2660)				2665	5				2670)	
Arg	Ası	o Ile	e Arg	g Thr	Glu	Glu	ı Ser	· I1e	Tyr	Gln	Ala	Cys	Ser	Leu	Pro
		267	5				2680)				2685	5		
G1u	Glu	ı Ala	a Ar	g Thr	Ala	ı Ile	His	s Ser	Leu	Thr	Glu	Arg	g Leu	ı Tyr	· Val
	2690	0				2695	5				2700)			
G1y	G1;	y Pr	o Me	t Phe	e Ası	ı Sei	Lys	s Gly	/ Glr	s Sei	Cys	G13	7 Туз	. Arg	g Arg
270)5				2710)				2715	5				2720

Cys	Arg	Ala	Ser	Gly	Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr
			2	2725				2	730				2	735	
Cys	Tyr	Val	Lys	Ala	Leu	Ala	Ala	Cys	Lys	Ala-	Ala	Gly	Ile	Ile	Ala
		2	2740				2	2745				2	2750		
Pro	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ile	Ser	Glu	Ser
-	2	2755				2	2760				2	2765			
G1n	G1ÿ	Thr	Glu	Glu	Asp	Glü	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
:	2770			•		2 7 75				2	2780		•		
Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr
278	5			- 2	2790	•	:	*. •		2795		٠.	:	. 2	2800
Asp	Leu	Glu	Leu	Ile	Thr	Ser	Cys	Ser	Ser	Asn	Val	Ser	Val	Ala	Leu
				2805				2	2810				· į	2815	
Gly	Pro	G1n	Gly	Arg	Arg	Arg	. Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr
			2820					2825					2830		
Ser	Ile	Ala	Arg	Ala	Ala	Trp	Glu	Thr	Val	Arg	His	Ser	Pro	Val	Asn
		2835					2840				•	2845			
Ser	Trp	Leu	Gly	Asn	Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg
	2850)				2855	5				2860)			
Met	. Val	Leu	ı Met	Thr	His	s Phe	Phe	Ser	Ιlε	Leu	Met	Ala	Gln	Asp	Thr
286	55		•		2870)				2875	;				2880
Let	ı Asp	Glr	n Asr	ı Leu	Asr	n Phe	e Glu	ı Met	Tyr	Gly	Ser	Val	Tyr	Ser	Val
				2885	;				2890)				2895	,
Sei	r Pro	Let	ı Asp	Leu	Pro	o Ala	a Ile	e Ile	Glu	ı Arg	, Leι	ı His	s Gly	Leu	Asp
			2900)				2905	5				2910)	
Ala	a Phe	e Se	r Lei	ı His	Th	r Ty	r Thi	r Pro	Hi	s Glu	ı Let	ı Thi	c Arg	y Val	. Ala
		291	5				2920	0				292	5		
Se	r Ala	a Le	u Ar	g Lys	Le	u G1;	y Ala	a Pro	Pr	o Lei	ı Arı	g Ala	a Trp	b Lys	Ser
	293	0				293	5				294	0			
Ar	g Al	a Ar	g Al	a Val	l Ar	g Al	a Se	r Leı	ı I1	e Se	r Ar	g Gl	y G1	y Arg	g Ala
								, 9:	2/11	7					

2945 2950 2955 2960

Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu 2965 2970 2975

Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp 2980 2985 2990

Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg 2995 3000 3005

Ala Arg Pro Arg Leu Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly 3010 3015 3020

Val Gly Leu Phe Leu Leu Pro Ala Arg 3025 3030

<210> 7

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 7

accugcccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60 cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacuggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360

ucaaagaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggggccc cgguucuuuu 540 ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780 ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200 ссисисссис ссесссски аасдинасив десдаадсед синддаанаа ддесддидид 1260 cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggeee ugucuucuug acgageauuc cuaggggucu uuccceucue geeaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 geggecaaaa gecaegugua uaagauaeae eugeaaagge ggeacaaeee eagugecaeg 1560 uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugeuuuae auguguuuag uegagguuaa aaaaaegueu aggeeeeeg aaceaegggg 1740 acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100

cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160 ccgagaccca uuucgaccuu gaaggggucc ucgggggggc cggugcucug cccuaggggc 2220 cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280 uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340 acgccaccgg cugugcccca gaccuaucag gucggguacu ugcaugcucc aacuggcagu 2400 ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460 aaccccucgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520 aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580 acauauggca aauuucucgc cgaugggggc ugcgcuagcg gcgccuauga caucaucaua 2640 ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700 caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccggguca 2760 gugacaaccc cccaucccga uauagaaaag guaggccucg ggcgggaggg ugagaucccc 2820 uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880 cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940 guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000 gucgccaccg acgcccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060 aauguagegg ucaeceaage uguegaeuue ageeuggaee eeaeeuueae uauaaeeaea 3120 cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180 agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240 guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300 accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360 cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucuc 3420 caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480 gccagagcca aggccccucc cccguccugg gacgccaugu ggaagugccu ggcccgacuc 3540 aagccuacgc uugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600 gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660 gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720 cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780 gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840

gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900 uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau caugggaggc 4140 ugguuagcgu cccagaucgc accacccgcg ggggccaccg gcuuugucgu caguggccug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg cgcgguccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620 uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680 cugaccucua aauuguuccc caagcugccc..ggccuccccu ucaucucuug ucaaaagggg 4740 uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800 aucucuggca auguccgccu gggcucuaug aggaucacag ggccuaaaac cugcaugaac 4860 accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920 cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980 cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040 caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100 cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160 gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220 cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280 ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgcu gcgggccacc 5340 ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400 ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460 auggeegagg aagagagega eeuugageee ucaauaceau eggagugeau geueeeeagg 5520 ageggguuue caegggeeuu aeeggeuugg geaeggeeug aeuaeaaeee geegeuegug 5580

gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucccccc 5640 cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760 ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacguccccu 5820 ggugagccgg cccccucaga gacagguucc gccuccucua ugcccccccu cgagggggag 5880 ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940 gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000 accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060 gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaaguge ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240 aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggecug cucceugece gaggaggece geacugecau acaeuegeug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugee gegeeagegg ggugeuaace acuageaugg guaacaceau caeaugeuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguaugc 6960 ggcaaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020 agagecuuca eggaggecau gaceagguae ueugeceeue euggugauee eeceagaeeg 7080 gaauaugacc uggagcuaau aacauccugu uccucaaaug ugucuguggc guugggcccg 7140 cggggccgcc gcagauacua ccugaccaga gacccaacca cuccacucgc ccgggcugcc 7200 ugggaaacag uuagacacuc cccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260 ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320

gacaccung accagaaccu caacuungag auguanggan caguanacuc cgugaanccu 7380
uuggaccuuc cagccanaan ugagagguna cacgggcung acgccunuuc naugcacaca 7440
uacucucacc acgaacugac gcggguggcu ucagcccuca gaaaacuung ggcgccaccc 7500
cucagggugu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560
aaagcggccg uungcggccg anaucucuuc aanunggcgg ugaagaccaa gcucaaacuc 7620
acuccanugc cggaggcgcg ccuacuggac uunuccagun gguucaccgu cggcgccggc 7680
gggggcgaca uunuucacag cgugucgcgc gcccgacccc gcucaunacu cuucggccua 7740
cuccuacuun ucgnaggggn aggccucuuc cuacucccg cucgguagag cggcacacac 7800
uagguacacu ccanagcnaa cuguuccun uunuunuun uunuunuun uunuunuun 7860
uunuunuunuu cuunuunuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980
augacugcag agagugccgu aacuggucuc ucugcagauc augu cgugaagguc cgugagccg 7980
augacugcag agagugccgu aacuggucuc ucugcagauc augu cgguaagguc cgugagccgc 7980

<210> 8

(211) 7994

<212> RNA

<213> Artificial Sequence

<220>

(223) Description of Artificial Sequence: replicon

<400> 8

accugcccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60 cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacugguu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 cgcuugcgag ugccccggga ggucucguag accgugcacc augagcaca auccuaaacc 360

ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggcugc ucugaugccg ccguguuccg gcugucagcg cagggggggcc cgguucuuuu 540 ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggeuggee acgaeggeg uuccuugege ageugugeue gaeguuguea eugaageggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugcegag aaaguaucca ucauggeuga ugcaaugegg eggeugeaua egeuugauce 780 ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200 ccucucccue-ccecccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260 cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggece ugucuucuug acgageauuc cuaggggucu uuccecucuc gecaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 geggecaaaa gecaegugua uaagauaeae eugeaaagge ggeacaaeee eagugecaeg 1560 uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740 acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggeegggg aagueeaaau eeugueeaca gueucueagu eeuueeuegg aacaaceaue 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100

	cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
	ccgagaccca	uuucgaccuu	gaaggggucc	ucgggggggc	cggugcucug	cccuaggggc	2220
	cacgucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
	uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
	acgccaccgg	cugugcccca	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
	ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
	aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggc	acauggcauc	2520
	aaucccaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
	acauauggca	aauuucucgc	cgaugggggc	ugcgcuagcg	gcgccuauga	caucaucaua	2640
	ugcgaugaau	gccacgcugu	ggaugcuacc	uccauucucg	gcaucggaac	gguccuugau	2700
	caagcagaga	cagccggggu	cagacuaacu	gugcuggcua	cggccacacc	ccccggguca	2760
	gugacaaccc	cccaucccga	uauagaagag	guaggccucg	ggcgggaggg	ugagaucccc	2820
	uucuauggga	gggcgauucc	ccuauccugc	aucaagggag	ggagacaccu	gauuuucugc	2880
	cacucaaaga	aaaaguguga	cgagcucgcg	gcggcccuuc	ggggcauggg	cuugaaugcc	2940
	guggcauacu	auagaggguu	ggacgucucc	auaauaccag.	.cucagggaga	uguggugguc	3000
	guegeeaceg	acgcccucau	gacgggguac	acuggagacu	uugacuccgu	gaucgacugc	3060
	aauguagcgg	ucacccaago	ugucgacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
	cagacuguco	cacaagacgo	ugucucacgo	agucagcgcc	gcgggcgcac	agguagagga	3180
•	agacagggca	cuuauaggua	a uguuuccacu	ggugaacgag	ccucaggaau	guuugacagu	3240
			a cgacgcaggg				
	accaccguca	ggcuuagag	c guauuucaac	acgcccggcc	uacccgugug	g ucaagaccau	3360
	cuugaauuuu	ı gggaggcag	u uuucaccggo	cucacacaca	uagacgccca	cuuccucucc	3420
	caaacaaago	aagcggggg	a gaacuucgcg	uaccuaguag	ccuaccaago	uacggugugc	3480
	gccagagcca	a aggccccuc	c cccguccugg	gacgccaugu	ggaagugcci	ı ggcccgacuc	3540
	aagccuacg	c uugcgggcc	c cacaccucuo	cuguaccguu	ı ugggcccuai	ı uaccaaugag	3600
	gucacccuc	a cacacccug	g gacgaaguad	aucgccacat	ı gcaugcaago	c ugaccuugag	3660
	gucaugacc	a gcacguggg	u ccuagcugga	a ggaguccugg	g cagccgucg	c cgcauauugo	3720
	cuggcgacu	g gaugcguuu	c caucaucgg	cgcuugcac	g ucaaccago	g agucgucguu	3780
	gcgccggau	a aggaggucc	u guaugaggc	ı uuugaugaga	a uggaggaau	g cgccucuagg	3840
					•		

gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900 uugcugcage aggecucuaa geaggeecag gacauacaae eegcuaugca ggeuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 geegeecuca ecagueeguu guegaeeagu accaecauce uucucaacau caugggagge 4140 ugguuagcgu cccagaucgc accacccgcg ggggccaccg gcuuugucgu caguggccug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg cgcgguccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620 uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680 cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740 uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800 aucucuggca auguccgccu gggcucuaug aggaucacag ggccuaaaac cugcaugaac 4860 accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920 cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980 cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040 caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100 cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160 gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220 cuaacagauc cgcccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280 ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgcu gcgggccacc 5340 ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400 ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460 auggeegagg aagagagega eeuugageee ucaauaceau eggagugeau geueeeeagg 5520 agcggguuuc cacgggcuu accggcuugg gcacggccug acuacaaccc gccgcucgug 5580

gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucccccc 5640 cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760 ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacguccccu 5820 ggugagccgg cccccucaga gacagguucc gccuccucua ugcccccccu cgagggggag 5880 ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940 gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000 accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060 gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaaguge ucgaegeeca uuaugaeuca gueuuaaagg acaucaageu ageggeuuce 6240 aaggucageg caaggeuccu caccuuggag gaggegugee aguugacuce accecauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480. gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggecug cucccugece gaggaggece geacugecau acaeuegeug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugee gegeeagegg ggugeuaace acuageaugg guaacaceau cacaugeuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugececue cuggugauce ecceagaceg gaauaugace uggageuaau aacauceugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140 gacccaacca cuccacucge cegggeugee ugggaaacag uuagacacue eccuaucaau 7200 ucauggcugg gaaacaucau ccaguaugcu ccaaccauau ggguucgcau gguccuaaug 7260 acacacuucu ucuccauucu caugguccaa gacacccugg accagaaccu caacuuugag 7320

<210> 9 · · · · · · · ·

<211> 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 9

accugcccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60 cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacuggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 cgcuugcgag ugccccggga ggucucguag accgugcacc 340

<210> 10

⟨211⟩ 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 10

accegeccu aauaggggcg acacucegec augaaucacu eceeugugag gaacuacugu 60 cuucacgcag aaagegucua geeauggegu uaguaugagu gueguacage eueeaggeec 120 ceeeuceeg ggagageeau aguggucuge ggaaceggug aguacacegg aauugeeggg 180 aagacuggu eeuuucuugg auaaaceeae ueuaugeeeg geeauuuggg egugeeeeg 240 caagacugeu ageegaguag eguuggguug egaaaggeeu ugugguacug eeugauaggg 300 ugeuugegag ugeeeegga ggueueguag acegugeaee 340

<210> 11

<211> 236

<212> RNA

<213> Artificial Sequence

⟨220⟩

<223> Description of Artificial Sequence: synthetic RNA

<400> 11

ageggeacae acuagguaca cuccauageu aacuguuccu uuuuuuuuuu uuuuuuuu 60

<210> 12

<211> 232

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 12

<210> 13

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 13

cgggagagcc atagtgg

17

<210> 14

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

 $\langle 223 \rangle$ Description of Artificial Sequence: synthetic DNA

<400> 14

agtaccacaa ggcctttcg

19

<210> 15

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 15

ctgcggaacc ggtgagtaca c

21

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 16

aacaagatgg attgcacgca

20

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220>-- -- · · · · ·

<223> Description of Artificial Sequence: synthetic DNA

<400> 17

cgtcaagaag gcgatagaag

20

<210> 18

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

1	1	Λ	a	>	1	8
`	4	w	. ,	_	- 1	O

gcactctctg cagtcatgcg gctcacggac

30

<210> 19

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 19

cccctgtgag gaactactgt cttcacgc

28.

<210> 20

<211> 24

.<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 20

ccgggagagc catagtggtc tgcg

24

<210> 21

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 21

ccactcaaag aaaaagtgtg acgagctcgc

30

<210> 22

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 22

ggcttgggca cggcctga

18

<210> 23

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

/9935	Description	٥f	Artificial	Sequence:	synthetic	DNA
くととろう	Describiton	OT	VI PITICIAI	pedacuec.	0 3 11 0110 0 10	

<400> 23

gcggtgaaga ccaagctcaa actcactcca

30

<210> 24

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 24

agaacctgcg tgcaatccat c

21

<210> 25

<21:1> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 25

cccgtcatga gggcgtcggt ggc

23

PCT/JP2003/015038

<210> 26

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 26

accagcaacg gtgggcggtt ggtaatc

27

<210> 27

<211> 18

<212> DNA

5,

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

<400> 27

ggcacgcgac acgctgtg

18

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<220> <223> Description of Artificial Sequence: synthetic DNA <400> 28 30 agctagccgt gactagggct aagatggagc <210> 29 <211> 20 <212> DNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence:synthetic DNA (primer) <400> 29 20 aacaagatgg attgcacgca <210> 30 <211> 20 <212> DNA <213> Artificial Sequence <220>

<223> Description of Artificial Sequence:synthetic

DNA (primer)

<400> 30

cgtcaagaag gcgatagaag

WO 2004/104198

20

<210≥ 31

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:synthetic DNA

<400> 31

gcactctctg cagtcatgcg gctcacggac

30

<210> 32

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic DNA

⟨400⟩ 32

ccctgtgag gaactactgt cttcacgc

28

<210> 33

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 33

ccgggagagc catagtggtc tgcg

24

<210> 34

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 34

ccactcaaag aaaaagtgtg acgagctcgc

30

<210> 35

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:synthetic
 DNA(primer)

⟨400⟩ 35

ggcttgggca cggcctga

18

<210> 36

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 36

gcggtgaaga ccaagctcaa actcactcca.

30

<210> 37

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 37

agaacctgcg tgcaatccat c

21

<210> 38

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 38

cccgtcatga gggcgtcggt ggc

23

<210> 39

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

 $\langle 223 \rangle$ Description of Artificial Sequence::synthetic DNA

<400> 39

accagcaacg gtgggcggtt ggtaatc

27

<210> 40

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 40

ggaacgcgac acgctgtg

18

⟨210⟩ 41

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence::synthetic DNA

<400> 41

agctagccgt gactagggct aagatggagc

30

特許協力条約に基づく国際出願願書 原本(出願用) - 印刷日時 2003年11月25日(25.11.2003) 火曜日 10時04分41秒

VIII-5-1	不利にならない開示又は新規 性喪失の例外に関する申立て 不利にならない開示又は新規性 喪失の例外に関する申立て(規 則4.17(v)及び51の2.1(a)(v))	本国際出願に関し、
	•	東レ株式会社は、 本国際出願の請求項に記載された対象が以下のよう に開示されたことを申し立てる。
VIII-5-1	開示の種類	刊行物
(i) VIII-5-1	 開示の日付:	2002年11月25日(25. 11. 2002)
(ii) VIII-5-1	開示の名称:	第25回日本分子生物学会年会プログラム・講演要旨
(iii)		集P386, W3aF-2
VIII-5-1	開示の場所:	
(iv)_		
VIII-5-1	本申立ては、次の指定国のため	すべての指定国
(v)	になされたものである。:	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/15038

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl ⁷ Cl2N15/11, Cl2N5/00, Cl2Q1/06, Cl2N7/00					
According to	o International Patent Classification (IPC) or to both nat	ional classification and IPC			
	S SEARCHED				
	ocumentation searched (classification system followed b C1 ⁷ C12N15/11, C12N5/00, C12Q1				
			in the Solds seembed		
	tion searched other than minimum documentation to the				
	lata base consulted during the international search (name bank/EMBL/DDBJ/GeneSeq, CAPLUS/N				
C. DOCU	MENTS CONSIDERED TO BE RELEVANT				
Category*	Citation of document, with indication, where app	propriate, of the relevant passages	Relevant to claim No.		
х	Wakita, T. et al., "Idenshigata 2a no C-gata 1-21 Kan'en Virus RNA Replicon no Juritsu", Dai 25 Kai The Molecular Biology Society of Japan Nenkai Program Koen Yoshishu, 25 November, 2002 (25.11.02), page 386, abstract.No.W3aF-2				
Y	EP 1043399 A2 (BARTENSCHLAGE 11 October, 2003 (11.10.03), Claims & CA 2303526 A1 & DE & AU 2518000 A & JP	1-21			
Further documents are listed in the continuation of Box C. See patent family annex.					
* Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed Date of the actual completion of the international search O2 February, 2004 (02.02.04) "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention cannot be considered novel or cannot be considered novel or cannot be considered novel or cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art document member of the same patent family Date of mailing of the international search report 17 February, 2004 (17.02.04)					
	Name and mailing address of the ISA/ Authorized officer				
	anese Patent Office	Telephone No.			

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/15038

		7
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	IKEDA, M. et al., "Selectable subgenomic and genome-length dicistronic RNAs deprived from an infectious molecular clone of the HCV-N strain of hepatitis C virus replicate efficiently i cultured Huh7 cells", J.Virol., (2002 March), Vol.76, No.6, pages 2997 to 3006, Fig. 8	1-21
. Y	Friebe, P. et al., "Sequences in the 5' non-translated region of hepatitis C virus required for RNA replication", J. Virol., (2001), Vol.75, No.24, pages 12047 to 12057, full text	1-21
Y	Lohmann, V. et al., "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line", Sceince, (1999), Vol.285, pages 110 to 113, Fig. 1	1-21
Y	WO 00/75338 A2 (THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICE), 14 December, 2000 (14.12.00), Claims; examples & EP 1185664 A2	1-21
Y	JP 2002-171978 A (Zaidan Hojin Tokyo-To Igaku Kenkyu Kiko), 18 June, 2002 (18.06.02), Claim 5; SEQ.ID.No.1 (Family: none)	1-21
	·	

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

国際調査報告	国際出願番号 PCT/JP03/15038			
A. 発明の属する分野の分類(国際特許分類(IPC)) Int. Cl' Cl2N15/11, Cl2N5/00,	C12Q1/06, C12N7/00			
B. 調査を行った分野 調査を行った最小限資料(国際特許分類(IPC)) Int. Cl ⁷ Cl2N15/11, Cl2N5/00,	C12Q1/06, C12N7/00			
最小限資料以外の資料で調査を行った分野に含まれるもの				
国際調査で使用した電子データベース(データベースの名称、 Genbank/EMBL/DDBJ/GeneSeq CAPLUS/MEDLINE/BIOSIS/WPII				
C. 関連すると認められる文献				
引用文献の カテゴリー* 引用文献名 及び一部の箇所が関連すると	関連する 関連する は、その関連する箇所の表示 - 請求の範囲の番号			
X Wakita, T. et al., "遺伝子型2 aのリコンの樹立"第25回日本分子生物旨集, 2002.11.25, p.386, Abst. No.	C型肝炎ウイルスRNAレプ 1-21 学会年会プログラム・講演要			
Y EP 1043399 A2 (BARTENSCHLAGER, R.) 請式の範囲参照 & CA 2303526 A1 & DE 19915178 A & JP 2001-017187 A				
C欄の続きにも文献が列挙されている。				
* 引用文献のカテゴリー 「A」特に関連のある文献ではなく、一般的技術水準を示すもの 「E」国際出願目前の出願または特許であるが、国際出願日以後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの「Y」特に関連のある文献であって、当該文献と他の1以上の文献(理由を付す) 「O」口頭による開示、使用、展示等に言及する文献「P」国際出願日前で、かつ優先権の主張の基礎となる出願「是際等の本地たの発送日				
国際調査を完了した日 02.02.2004	国際調査報告の発送日 17.2.2004			
国際調査機関の名称及びあて先日本国特許庁(ISA/JP)	特許庁審査官(権限のある職員) 4B 9639 新留 豊			
郵便番号100-8915 東京都千代田区額が関三丁目4巻3号	 電話番号 03-3581-1101 内線 3448			

国際調査報告

<u> </u>	四次的五十六日		
C(続き).	関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連すると	きは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	Ikeda, M. et al., "Selectable subger dicistronic RNAs deprived from an in of the HCV-N strain of hepatitis C efficiently i cultured Huh7 cells" Vol. 76, No. 6, pp. 2997-3006, Fig. 8	nomic and genome-length infectious molecular clone virus replicate J. Virol., (2002 Mar),	1–21
Y	Friebe, P. et al., "Sequences in the of hepatitis C virus required for RIJ. Virol., (2001), Vol. 75, No. 24, p	NA replication"	1-21
Y	Lohmann, V. et al., "Replication of virus RNAs in a hepatoma cell line" Vol. 285, pp. 110-113, Fig. 1 参照		1-21
Y .	WO 00/75338 A2 (THE GOVERNMENT OF TAMERICA AS REPRESENTED BY THE SECRE HEALTH AND HUMAN SERVICES), 2000.12 請求の範囲,実施例参照 & EP 1185664 A2	TARY, DEPARTMENT OF	1-21
Υ	JP 2002-171978 A (財団法人 東京都區 請求の範囲5, SEQ ID NO.1 参照 (ファ		1-21
	,		